ESSAYS ON HEALTH ECONOMICS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Eamon Joseph Molloy

August 2012

ESSAYS ON HEALTH ECONOMICS

Eamon Joseph Molloy, Ph.D.

Cornell University 2012

The essays of this dissertation study the effect of alcohol advertising on individual drinking, alcohol firm advertising decisions, and the relationship between education and mortality.

The first essay focuses on the possible effects of alcohol advertising on youth drinking. Researchers still disagree about how advertising affects alcohol consumption. This disagreement largely arises because alcohol firms target marketing at people who already drink. Drinkers prefer particular media; firms recognize this and target alcohol advertising at these media. Endogenous targeting of alcohol advertisements presents a challenge for empirically identifying a causal effect of advertising on drinking. In this chapter, I overcome these challenges by leveraging a plausibly exogenous source of variation in advertising exposure, and by utilizing novel data with detailed individual measures of media viewing and alcohol consumption. I adopt three approaches to control for endogeneity bias due to targeting. First, I use average audience characteristics of the media an individual views to capture targeting. Second, I use media fixed effects to directly control for media choice. Third, I exploit variation in advertising exposure due to a 2003 change in an industry-wide rule that governs where firms may advertise. I use the rule change as an instrument for exposure to alcohol advertising. Though the unconditional correlation between advertising and drinking is strong, this relationship is not robust to more rigorous controls for targeting and to the use of an instrumental variables estimator. The results suggest that any effect of alcohol advertising on youth drinking is modest.

The second essay studies the effects of the end of the liquor broadcast advertising ban

on firm behavior. I study which firms and brands first took advantage of this new medium I study which spirits brands take advantage of the newly available medium of television. I compare the consumer characteristics and market competition of brands that transition to television advertising to those that do not, using two different definitions of television advertising adoption. I model brand-level, yearly television advertising spending and estimate hazard models of the transition to the use of television advertising. I find evidence that competitive pressure correlates with a brand's adoption of the "new" medium. Firms that are dominant in their market are much more likely to adopt television advertising when their competitors possess a larger share of the market. However, I find little evidence that the demographic characteristics (age, gender, race, income, education, magazine reading, and television viewing) and alcohol consumption of a brand's consumers are related to the adoption of television advertising. The results suggest that television advertising in the spirits market may play larger role dividing market shares than growing market size.

The third essay revisits the question of whether people live longer if they get more education or if people who get more education have unobservable traits and habits that cause them to live longer. Like previous studies, we use compulsory schooling laws as instruments for education, However, we use better instruments and Panel Study of Income Dynamics data that include each respondent's date and cause of death. We find our compulsory schooling instruments are stronger predictors of education than those used in previous studies. However, relying on within-state variation greatly reduces the predictive power of our instruments, which only weakly predict educational attainment. We model three different measures of mortality: probit models of mortality over 5- and 10-year age spans and continuous-time survival models of the number of months a person lives past forty years of age. We confirm a strong statistical association between education and mortality in all three model types. However, due to the weakness of our instruments, our results are imprecise and provide little useful insight into whether education reduces mortality. We show the relationship between schooling and mortality is strongest for post-secondary education, though there exists little evidence in the literature concerning whether this link is causal.

BIOGRAPHICAL SKETCH

Eamon Joseph Molloy was born in Queens, New York. He graduated from Regis High School in 2001 and earned his Bachelor of Science degree in Policy Analysis and Management from Cornell University in 2005. After graduation, Eamon worked as a research assistant for Rosemary Avery, Donald Kenkel, Dean Lillard, and Alan Mathios in the Department of Policy Analysis and Management at Cornell University. In that role, Eamon supported a National Cancer Institute funded study on the effects of the advertising of tobacco and smoking cessation products on smoking behavior.

In 2007, Eamon joined the doctoral program in economics at Cornell University. He served as a teaching assistant for undergraduate classes on public finance and risk management. As a graduate research assistant, he aided in a National Cancer Institute funded study on the strategic marketing of cigarettes conducted by Dean Lillard and Andrew Sfekas. He also collaborated with Dr. Lillard and Dr. Sfekas to study the impact of cigarette prices and taxes on youth smoking initiation. Professor Donald Kenkel, Professor Michael Lovenheim, and Dean Alan Mathios supervised his dissertation research. Eamon graduated with his Doctor of Philosophy in economics in August 2012. Following graduation, he began work as an analyst at the Congressional Budget Office.

To My Parents

ACKNOWLEDGMENTS

I acknowledge and thank several funding sources that made this research possible: The Cornell Population Center, The Bronfenbrenner Life Course Center, The National Institute on Alcohol Abuse and Alcoholism, and the Department of Policy Analysis and Management of Cornell University.

I have benefited immensely from the assistance of my committee: Donald Kenkel, Alan Mathios, and Michael Lovenheim. Their expert feedback and support have been invaluable throughout the writing of this dissertation and, more generally, throughout my time in graduate school. I also owe a considerable debt to Dean Lillard for his mix of encouragement, insistence on the active voice, and patient mentoring.

The faculty and staff of Cornell University have also been an amazing source of knowledge and support. I would particularly like to thank John Abowd, John Cawley, George Jakubson, and Jordan Matsudaira for extremely valuable guidance on the chapters of this dissertation. I am thankful to Hua Wang for her advice, data expertise, and friendship. I also benefited greatly from the aid of Rosemary Avery and her team of excellent research assistants, particularly Jonathan Cantor, Rebecca Keable, and Mandela Steele-Dadzie. I thank Joshua Kim and Ben Ost for expert research assistance as well. I also greatly appreciate Eric Maroney and Geysa Smiljanic for their help.

My fellow graduate students have been an indispensable source of friendship and assistance. I thank all of the participants of the Policy Analysis and Management Research in Progress seminar for great feedback the numerous times I have presented this work there. I especially want to thank Kiel Albrecht, Catherine Maclean, Alex Rees-Jones, Kevin Roth, Matthew Sweeney, Michael Strain, Douglas Webber, and Chen Zhao for help both related to and unrelated to economics.

I thank Lindsay Gerolimatos, whose love and support were integral to the writing of this dissertation. Finally, I thank my mother, father, sister, and brother for being continuous sources of love and encouragement throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

THIS AD'S FOR YOU: TARGETING AND THE EFFECT OF ALCOHOL ADVERTISING ON YOUTH DRINKING

This research is supported by a Cornell Population Center Rapid Response Grant, the Bronfenbrenner Life Course Center Innovative Research Program, and award 1R01-AA019777-01A1 from the National Institute on Alcohol Abuse and Alcoholism.

1.1 Introduction

The consequences of alcohol abuse are serious and affect both under-age drinkers and drinkers of legal age. In 2004, 51.6 percent of youth ages 18 to 20 and 68.6 percent of youth ages 21 to 25 reported drinking in the past month (SAMSHA, 2007). Although many drink in moderation, 36.2 percent of 18- to 20-year-olds and 46.1 percent of 21- to 25-year-olds report having had five or more drinks in one occasion. This heavy drinking is a serious public health concern. The CDC estimates that in 2004 there were 75,766 alcohol-attributable deaths in the United States (CDC, 2004). Roughly 5,000 of those who die yearly are youth below the legal drinking age (Hingson and Kenkel, 2004). These deaths include approximately 1,900 vehicle accident fatalities, 1,200 accidents of other types, 1,600 homicides, and 300 suicides. Behind smoking and obesity, alcohol abuse is the third leading cause of preventable death in the U.S. (Mokdad et al., 2004). Carpenter and Dobkin (2009) find that a 10 percent increase in the number of drinking days leads to a 4.3 percent increase in mortality for 21-year-olds. Alcohol is associated with serious non-fatal consequences such as injuries, violence, and sexual assault. Hingson et al. (2005) find alcohol is implicated in approximately 70,000 sexual assaults and half a million injuries on college campuses annually. Bouchery et al. (2011) estimate that the economic cost of excessive drinking in the U.S., which is largely due to lost productivity, was $223.5 billion in 2006.

While alcohol abusers impose many serious costs on themselves, their drinking also imposes significant costs on others. Alcohol-related traffic accidents, crime, sexual assault, and risky sex create negative externalities. Kenkel (1993) finds that drunk drivers impose $17.6 billion (1986 dollars) in costs due to the injury and death of others. Carpenter (2007) shows that alcohol use causes 18- to 20-year-olds to commit nuisance and property crime. Chesson, Harrison and Kassler (2000) estimates that alcohol-attributable sexually transmitted diseases create externalities of at least $556 million annually. Drinkers do not fully internalize these and other costs they impose by consuming alcohol irresponsibly. This market failure creates a clear rationale for government intervention to curb the social costs associated with alcohol.

The external costs associated with youth alcohol abuse and the large sums firms spend marketing alcohol create concern that advertising may cause youth drinking. To address this concern, the alcohol industry regulates its marketing to avoid appealing to youth and to promote responsible drinking. However, large advertising expenditures make youth exposure inevitable. In 2005, the alcohol industry spent $1.10 billion on 292,944 television advertisements and $429 million for 4,677 magazine advertisements (Kantar Media). Though alcohol firms are subject to standard truth-in-advertising regulations, they face no industry-specific federal advertising regulations. The federal government simply monitors whether firms follow their own rules. State governments legislate rules affecting less prevalent and more local types of advertising, such as on billboards and at points of sale. Despite the lack of legal restrictions, trade associations representing the wine, beer, and distilled spirits industries establish and enforce rules intended to reduce the impact of advertising on youth. All three trade associations require that alcohol advertisements only appear in media in which more than 70 percent of the audience is of legal drinking age. Despite these efforts, the average youth saw 41 percent more television alcohol ads in 2005 than in 2001 (CAMY, 2010b).

The alcohol industry maintains that advertising influences the brands consumers choose but does not affect whether and how much they drink (Beer Institute, 2011*b*). If this claim is true, alcohol advertising does not threaten public health. That is, if non-drinkers are not being persuaded to drink and drinkers are not being persuaded to drink more, then alcohol advertising is not responsible for any of the negative externalities associated with alcohol abuse. Establishing the veracity of this claim is of critical interest to government regulators, policymakers, and academic economists. In this paper I test this claim and focus on the basic empirical question of whether advertising affects alcohol consumption.

This question has been studied prior to this paper. A review by Hastings et al. (2005) concludes that there is convincing evidence from numerous studies that advertising has an effect on alcohol consumption. However, a review by Nelson (2010) concludes that the empirical methods employed by previous studies do not sufficiently control for endogeneity. My own review of the literature confirms Nelson (2010). As a consequence, the current estimates of the effect of alcohol advertising on consumption are not credible and need to be improved.

Researchers disagree over whether advertising causes people to drink largely because existing studies do not control for the multiple ways in which advertising exposure could be endogenous. Individuals choose which media to view, and people who drink may systematically select into viewing television programs and magazines which non-drinkers will not choose to view. Alcohol firms may target advertisements at those programs and magazines with more viewers who drink in order to compete for market share among drinkers. If they do so, exposure to alcohol advertising will be strongly and positively correlated with the drinking of viewers of those targeted media. However, this association reflects a reverse causality: People do not drink because they see more advertising; rather, drinkers see more advertisements because they drink. Any credible estimate of the effect of advertising must account for this targeting.

This paper addresses targeting using three different empirical methods, each of which represents a significant improvement over the current methods used in the literature: the first method controls for media-specific average audience demographic characteristics and drinking behavior, the second estimates the effect using within-media variation, and the third exploits quasi-exogenous variation with a "natural experiment" methodology.

First, I control for the average audience demographic characteristics and drinking behavior of the media which an individual views. If firms select which magazines and programs to advertise on based on the composition of their audiences, people who view magazines and programs with audiences coveted by alcohol firms will see more advertising. However, choosing programs popular with alcohol advertisers may also be correlated with your own drinking. For example, football games may be popular with beer advertisers because young men who drink beer watch them. In this method, I control for the fact that, on average, you prefer shows popular with young men and beer drinkers and that this preference may be related to both your exposure to alcohol advertising and your drinking. This approach controls for the very characteristics that an alcohol firm likely considers when selecting where to target advertisements.

Second, I directly control for the media individuals view that may lead them to see more alcohol advertising. I employ media choice controls that directly account for the average drinking behavior of each television program and magazine in my data. In these models, I control for the fraction of issues or television program airings a person reports viewing for each magazine and television program in my survey data. I then use only variation in advertising within a program or magazine over time to estimate the effect of advertising. This method requires fewer assumptions about the targeting behavior of firms than the first method and uses decidedly stronger controls for the endogeneity of media choice than previous work.

Third, I instrument for exposure using a 2003 change to the self-regulation of alcohol advertising that introduced exogenous variation in advertising exposure on several television programs and magazines. The variation I use in this method is likely driven by the change in regulation and plausibly uncontaminated by targeting. This method allows me to address the concern that firms may target advertising based on program characteristics that change over time. This approach also directly evaluates the effect on youth drinking of a policy meant to reduce youth exposure to alcohol advertising. The estimates generated by this "natural experiment" are both uncontaminated by bias from targeting and directly relevant to current and future policy.

Each of these methods confronts the reverse causality targeting may cause, and each represents an improvement over existing studies. Consequently, this study more credibly estimates the effect of advertising on drinking than any previous study.

In addition to offering improved methods, this paper also offers a significantly improved measure of alcohol exposure. I use novel, individual-level data that combine detailed alcohol consumption data and detailed data measuring media consumption. I combine the latter data with comprehensive data on alcohol advertisements on television and in print. The result is a measure of advertising exposure that varies across individuals, time, and geography and is more comprehensive in the number of media it includes than previous measures. This exposure measure is superior to any measure in the published literature.

Though my models without controls for targeting, which are similar to those in previous studies, confirm a strong, positive relationship between alcohol advertising and youth (ages 18 to 24) drinking, these results are sensitive to controls for firm advertising targeting. This positive, statistically significant relationship between drinking and advertising is robust to controlling for average media audience characteristics. These results imply effects of more modest economic significance than previous studies; a halving of average magazine and television advertising exposure would reduce under-age youth drinking by 2.0 percentage points and over-age drinking by 2.2 percentage points. The size of these effects are conservative upper-bounds, as estimates only using media average controls may still suffer from targeting bias. More credible estimates, which directly control for media choice, find a much smaller and statistically insignificant effect: A halving of average advertising exposure would increase under-age and over-age drinking prevalence by 0.2 percentage points, reduce consumption among drinkers by 1.0 drinks per month for the under-age and 1.2 drinks per month for the over-age, and *increase* under-age "heavy"[1] drinking by 0.2 percentage points and decrease over-age "heavy" drinking by 0.1 percentage points. Although estimated with less precision, these results and those from instrumental variables models indicate no general statistically significant effect of alcohol advertising on youth drinking.

The rest of the paper is organized as follows: Section 2 reviews several theories concerning advertising's role in influencing demand and the accompanying empirical evidence. I also discuss the alcohol industry and its self-regulation of advertising. Sections 3 and 4 introduce my data and econometric approach. Section 5 reports my model results. I discuss my results and conclude in Section 6.

1.2 Background

Advertising and Demand

While advertising is ubiquitous in modern life, economic theory has struggled to explain its existence. Standard consumer theory does not predict that Clydesdale horses or cartoon frogs will affect a rational and well-informed individual's consumption of Budweiser. There are three main views that attempt to explain why consumers respond to advertising (Bagwell, 2007). First, the "persuasive view" of advertising regards it as altering consumers' tastes, convincing them of the superiority of one of several identical goods. This shift in

[1] I define "heavy" drinking as 33 or more drinks per month.

preferences leads advertising firms to face more inelastic demand for their products, allowing them to charge higher prices. However, most economic models do not assume preferences are so malleable. Second, rather than alter preferences, the "informative view" theorizes that advertising provides consumers with a convenient source of product information. Studying the role of product information in consumer behavior, Nelson (1970) distinguishes between the role of advertising in the markets for search goods, whose quality can be determined prior to consumption, and experience goods, whose quality must be experienced through consumption. More specifically, Nelson (1974) contends that advertising of experience goods provides indirect product information that signals quality. Given the importance of its taste, alcohol consumption seemingly has a significant "experience" component. Third, the "complementary view" posits that the advertising and consumption of a product are complements. Stigler and Becker (1977) outline a model with stable preferences but a complementarity between advertising and consumption. The model implies that viewing the advertising of a product raises the marginal utility of consuming that product, leading to increased demand. One could think of advertising associating an alcohol brand with a particular "image" and consumers gaining added utility from drinking that brand and being associated with that "image." Each of these views suggests a different mechanism for how advertising might affect drinking, though they certainly are not mutually exclusive. Importantly, none of these theoretical links between advertising and consumption definitively predicts whether alcohol advertising increases alcohol consumption or simply alters brand choice.

Empirical research on the relationship between advertising and health behaviors, namely drinking and smoking, has produced mixed evidence of causal effects of advertising. A literature review of the effect of advertising on smoking concludes that there is no consensus on the effect of advertising on cigarette demand (Chaloupka and Warner, 2000). In a more critical review, Heckman, Flyer and Loughlin (2008, pg. 43) observe that the published evidence that purports to link cigarette advertising and smoking "emerge(s) from empirical implementations that fall far short of those required to establish well-founded causal relationships." However, Avery et al. (2007) demonstrate a causal effect of smoking cessation product advertising on quitting behavior. Research on the effect of alcohol advertising on consumption is inconclusive. A recent review of longitudinal studies of the effects of advertising on drinking and smoking finds that methodological shortcomings rule out causal inference (Nelson, 2010). In contrast, Hastings et al. (2005) concludes that there is compelling evidence for an effect of advertising on youth drinking. The contrasting opinions in these reviews represent a disciplinary divide regarding the effects of alcohol advertising: public health research generally concludes that advertising matters, while economics research typically is more skeptical of advertising's effects. This disagreement largely rests on how studies measure alcohol advertising exposure and how they confront the endogeneity of this exposure.

Numerous studies of alcohol advertising estimate the relationship between aggregate measures of advertising and total alcohol sales. Duffy (1991) and Nelson (1999) model demand systems for alcohol and find no association between national-level advertising expenditures and national alcohol sales in either the UK or the U.S. More extreme sources of variation in aggregate advertising levels are partial and full alcohol advertising bans. Nelson (2003) finds state bans on billboard advertising do not affect consumption. However, international studies show a significant reduction in aggregate alcohol consumption in OECD countries that instituted alcohol advertising bans (Saffer, 1991; Saffer and Dave, 2002). An extension of the aggregate advertising approach directly measures the impact of market-level advertising on drunk driving deaths. Saffer (1997) finds a relationship between market-level radio, television, and billboard advertising and market-level motor vehicle fatalities. Though studies of aggregate advertising and consumption provide evidence of the relationship between advertising and consumption, they largely do not control for the endogeneity of advertising and cannot detect advertising effects in specific subpopulations. In particular, these studies cannot isolate the effect of advertising on youth.

A more tailored approach uses survey data to study the association between alcohol advertising and individual-level alcohol consumption. These studies typically use one of three measures of advertising exposure: market-level exposure averages or advertising expenditures, self-reported alcohol advertising exposure, or individual-level exposure estimates. Saffer and Dave (2006) use the NLSY97 and Monitoring the Future surveys to show a link between the number of advertisements appearing in a youth's city and his past 30-day drinking and bingeing. However, this study does not entirely control for the targeting of advertising at specific media markets that may contribute to a positive correlation between market-level advertising and drinking. Additionally, because media choices are quite varied, the market-level average youth exposure to advertising is likely an inaccurate measure for any given individual's actual exposure. In studies that use the second measure, youth are asked to report how many alcohol advertisements they recall seeing in a given time period (typically one month). These studies all find a positive association between recalled advertising measures and youth drinking (Connolly et al., 1994; Wyllie, Zhang and Casswell, 1998*a*,*b*; Stacy et al., 2004; Snyder et al., 2006). However, youth who already drink (or are inclined to drink) may recall advertising better because they are interested in or aware of alcohol advertisements. Consequently, the measured correlation between advertising recall and drinking may be spurious. Furthermore, even if these measures are unbiased, they do not overcome targeting; youth who drink may consume media where advertising is more prevalent and subsequently may recall seeing more ads.

My study is most similar to research that uses the third measure, individual-level exposure estimates. These studies first ask youth to report their viewing of specific media (television programs and magazines), then they match the number of advertisements on each medium to each viewer of the medium and sum these measures across all media. Because respondents are not asked to specifically recall alcohol advertisements, these studies are not contaminated by recall bias.[2] Stacy et al. (2004), Ellickson et al. (2005), and Collins et al. (2007) use exposure estimates based on individual media viewing and show a significant relationship between these advertising measures and youth drinking. Some of these estimated relationships are large in magnitude: Stacy et al. (2004) finds a one standard deviation increase in viewing television programs with alcohol ads is associated with a 44 percent increase in beer use. The second and third exposure measures appear only in public health research. These studies do not account for targeting and other sources of endogeneity that may bias upward estimates of the effect of advertising.

This study confronts several issues in the extant literature. I use individual-level consumption data, allowing me to model the specific drinking behavior of youth. I also use individual-level measures that estimate exposure based on the magazines and televisions programs youth actually report viewing. These measures are free of obvious recall bias and more likely predict the advertising an individual saw. Finally and most importantly, I address the endogeneity of media choice. Youth do not choose magazines and television programs at random. Similarly, alcohol firms do not aimlessly choose where to advertise. The critique leveled in Heckman, Flyer and Loughlin (2008, pg. 42) concerning the state of cigarette advertising research is just as valid for alcohol: "Specifically, these studies ignore the consequences of human choice for the validity of their statistical analyses." The targeting of advertising at media popular with drinkers and at media markets populated with drinkers creates a serious threat to causal inference. This study uses several methods to address this threat. Each of my models includes market fixed effects to control for geographic drinking heterogeneity. I control both for the general characteristics of the media one consumes and directly for the media he chooses. Finally, I use a change in industry advertising self-regulation that is plausibly unrelated to individual media choices, as a "natural" experiment.

[2] This argument assumes youth do not remember viewing media because there were alcohol advertisements on them.

The novel data and methods used by this study provide more credible estimates of the effect of advertising on drinking.

Alcohol Industry Self-Regulation of Advertising

While the sale of alcohol in the U.S. is subject to significant government regulation, the advertising of alcohol is largely self-regulated.[3] The history of this self-regulation dates back to at least the end of prohibition. Months after the passage of the 21st Amendment re-legalizing alcohol, liquor firms agreed not to advertise on radio and subsequently extended this agreement to include television in 1948 (Frank, 2008). Facing decades of declining sales, this agreement ended in 1996, and liquor advertising began appearing in broadcast media. The large increases in liquor advertising in the early 2000's drives much of the temporal variation in advertising used in this study. However, because liquor firms likely choose to place these liquor advertisements in media popular with drinkers, targeting is still a concern.

Industry groups representing beer, wine, and liquor producers each outline a "code of responsible practice" for advertising. These codes particularly focus on avoiding marketing to youth. Rules prohibit certain advertising content (e.g. ads depicting Santa Claus or cartoon characters) and specify guidelines for advertising placement to limit the number of youth viewing an advertisement (Beer Institute, 2011; DISCUS, 2011; Wine Institute, 2011). In 2003 the Beer Institute and Distilled Spirits Council of the U.S. (DISCUS) moved from a voluntary rule of only placing ads in media where at least 50 percent of the audience is of legal purchasing age (LPA) to a 70 percent rule. The Wine Institute moved to this stricter standard in 2000 (FTC, 2003). The Federal Trade Commission (FTC) recommended this shift in a 1999 report reviewing industry advertising (FTC, 1999).[4] The 70 percent guideline

[3]Though the advertising of alcohol is subject to significant truth-in-advertising rules, unlike tobacco products, there are no specific federal regulations on advertising.

[4]Though the FTC does enforce regulations concerning alcohol advertising, the Commission monitors alcohol advertising on an ongoing basis. In 1998, Congress specifically requested the FTC investigate marketing

roughly corresponds to the percentage of the U.S. population age 21 or older. The stated purpose of this new rule is to avoid "over-targeting" youth with alcohol advertisements. More recently, in 2011 all three industry groups moved to a rule of 71.6 percent, following the 2010 U.S. Census measure of the percentage of the population that is legal purchasing age. This change also followed a 2008 FTC report suggesting the industry use Census data to determine advertisement placement rules (FTC, 2008).

Along with general rules about advertising content and placement, the advertising codes contain specific instructions regarding code compliance that this study exploits. Each code specifies which data sources are appropriate measures of audience composition for a given media (television, magazine, radio, newspaper, and internet). The code requires that firms use recent data to measure audience composition; if possible firms should draw data from the previous two quarters. The code also compels firms to audit past advertisement placements to determine if they violated the code. If a firm determines advertisements in a particular medium violate the code, the code states the firm should remove future advertisements in that medium. The advertising code appears to have had some effect on the advertising decisions of alcohol firms. In 2002, prior to the 70 percent change, 2.6 percent of broadcast television network advertisements and 12.2 percent of cable network advertisements aired on programs with less than 70 percent LPA audiences (CAMY, 2010b). In 2005, no broadcast advertisements and 6.3 percent of cable advertisements violated the 70 percent LPA audience rule. Between 2002 and 2005, advertisements placed in magazines with less than 70 percent LPA audience also declined from 10.8 percent to 0.3 percent (CAMY, 2007).

In what follows, I use this rule change in an instrumental variables model to test for a causal effect of advertising on drinking. Using the recommended ratings data, I isolate the decline in advertisements in particular media (magazines and television programs) due to the change from a 50 percent LPA rule to the 70 percent rule. This variation in advertising

to youth and, if necessary, work with industry to develop more effective self-regulation.

is arguably unrelated to an individual's inclination to drink, making it an ideal source of variation in alcohol advertising exposure.

1.3 Data

Simmons National Consumer Survey (NCS)

The Simmons National Consumer Survey is uniquely useful for estimating advertising exposure and controlling for advertising targeting. The NCS includes extremely detailed questions on media viewing. Because of this detail, firms use the NCS to develop marketing plans. Given that marketers use the NCS to target advertisements at consumers, the survey is particularly well suited to study the effects of advertising. I can observe much of the same information alcohol firms use to choose programs and magazines for their advertisements, reducing concerns of omitted variables bias.

I use data from the 2000 to 2007 waves of the NCS. Administered biannually, the NCS uses a multi-stage stratified probability sample that is drawn from random digit dial sampling frames that exclude Hawaii and Alaska. The NCS collects detailed information on consumption of all types, including detailed information on the types and amount of alcohol a person consumed in the prior 30 days. Simmons collects a rich set of demographic characteristics on the individual (age, race, sex, religion, education, employment status, and schooling status) as well as household characteristics (income, number of adults, age of household head, and language spoken in the home). The NCS collects data on media consumption that is impressive in range, detail, and depth. For example, over these years the NCS asked whether respondents read each of 173 popular consumer magazines and viewed 2,510 television shows. For each magazine, the NCS asks respondents to indicate how many of the last four issues they have read. Respondents also indicate how often they watch each television program and what times of day they typically watch. Finally, all respondents (including 18 to 20 year olds) answer a section on alcohol consumption, listing the number of drinks consumed in the last month for each of several categories and brands (e.g., light domestic beer, rum, *Miller Lite*, *Bacardi*). The NCS yields a sample of all adults (age 18 and older) that, with sample weights, represents the US population living in households in the contiguous US. The pooled sample from the 2000 to 2007 waves consists of 174,253 adults. Of these, 7,505 are ages 18 to 20 and 6,634 are ages 21 to 24. I exclude 105 youth reporting drinking more than 150 drinks in the last 30 days due to data validity concerns. Appendix A describes the NCS in more detail and compares the NCS alcohol consumption data to an established national survey on substance abuse.

The comprehensive estimates of advertising exposure in this study include national broadcast and cable advertising and local broadcast (spot) advertising. To estimate exposure to locally airing television ads one must identify the Designated Marketing Area (DMA)[5] where an individual lives. The NCS identifies the DMA of respondents only for those who live in the 14 largest DMAs.[6] To identify the DMA in which other NCS respondents live, I use other geographic identifiers, including state of residence and the general size of the DMA, and a matching algorithm. The algorithm I developed allows me to match 78.1 percent of youth NCS respondents to a DMA. I cannot identify the DMA of the remaining 21.9 percent because these respondents live in states with multiple DMAs of roughly the same size. I only include respondents matched to a DMA in my analyses; The remaining age 18 to 24 sample includes 12,752 persons. Appendix A details the DMA matching data and algorithm.

[5]A DMA is a group of counties that share the same broadcast televisions stations, typically named for the largest city or cities in the area. Nielsen Media Research defines the boundaries of the 210 U.S. DMAs.

[6]Nielsen measures DMA size by the number of households with a television within the DMA.

Kantar Media Magazine and Television Advertising Data

I use data on advertising of alcohol that appeared in consumer magazines from 2000 to 2007. Kantar Media[7] compiled these data. For each magazine advertisement the data include: the alcohol brand and sponsor, the magazine, the issue (date), and the page size of the advertisement. Alcohol advertisements appeared in 147 of the 173 magazines included in the NCS during the sample period. I sum across all alcohol brands to generate the total number of alcohol advertisements that appeared in each magazine in each year. Figure 1.1 shows the general trends in the number of magazine alcohol advertisements from 2000 to 2007. The data show a decline in total magazine advertising, which is mostly driven by a reduction in magazine spirits advertising. I describe how I use these magazine-level data to generate individual exposure estimates below.

I also use Kantar Media data on television alcohol advertising from 2000 to 2007. These data include: the alcohol brand, the sponsor name, the name of the program during which it aired, when it aired (date and time), the coverage of the message (national broadcast TV, local broadcast TV, or cable TV), its length (in seconds), and the geographic media market where it aired. The television data track messages aired nationally or aired locally in 110 DMAs. The data represent the universe of television alcohol advertisements appearing nationally and locally on broadcast networks.[8]

Figure 1.2 displays general trends in monthly numbers of television alcohol advertisements in total and by alcohol type. Firms vary their advertising significantly over the calendar year, with peaks in advertising over summer and winter holiday months. The data also show a distinct, cross-year increase in advertising, largely driven by an increase in television

[7] Kantar Media was previously known as TNS Media and Competitive Media Reporting (CMR).

[8] Kantar does not measure local cable advertising and only measures cable advertising on 44 popular, national cable networks. Though my data only cover the 110 most populous of the 210 U.S. DMAs, the data include local broadcast advertising for every DMA I identify in the NCS.

spirits advertising.[9] Note that this time period includes a large increase and decline in the advertising of flavored malt beverages (e.g., *Smirnoff Ice* and *Mike's Hard Lemonade*). Given their sweet taste and bright packaging, the public health community was concerned that this new beverage type particularly appealed to youth (Mosher and Johnsson, 2005). Though I exploit temporal variation in advertising, I use wave fixed effects to control for any seasonality or larger time trends in drinking. Figure 1.3 summarizes these trends by splitting the data by advertising type. By simple count, a large majority of the advertisements appear on cable. However, broadcast programming typically attracts more viewers, making its advertisements more expensive. In 2005, alcohol firms spent roughly equal amounts on national cable and broadcast advertising.

Television and Magazine Ratings

I use magazine and television ratings data recommended by industry advertising codes to determine which programs are eligible for alcohol advertisements under the industry legal purchase age (LPA) audience rule. Nielsen Media Research provided me with television ratings data for 474 popular broadcast and cable television programs. I use these yearly ratings data to construct the LPA audience percentage for each program. Because firms can purchase daily Nielsen ratings for a program, I cannot know exactly which data firms use. However, these ratings estimates come from the same source alcohol firms must use to determine if an advertising placement complies with the code.

I use similar yearly LPA audience estimates for magazines, also based on the recommendations of industry advertising codes. I create these magazine ratings data by aggregating magazine reading data from the Simmons NCS. The codes list the NCS as one of two ac-

[9] Spirits producers appear to substitute television for magazine advertising following the end of their self-imposed ban on broadcast advertising.

ceptable sources of ratings data[10] to determine a magazine's LPA audience. I combine the NCS Adults survey, the same data described above, and the NCS Teens survey (ages 12 to 17), which also contains data on magazine reading. I then calculate the percentage of readers age 12 to 20 of each magazine using provided nationally representative weights. These percentages are precisely what the advertising code expects firms to use to comply with the 70 percent LPA rule.

Instrument Selection and Creation

I use the television and magazine ratings data to isolate variation in advertising in media directly affected by the change in the LPA audience rule. I isolate this variation by identifying programs and magazines that change "eligibility" for alcohol ads at some point in my sample period (2000 to 2007).[11] These magazines and television shows meet two criteria: the program or magazine was affected by the rule change (i.e., it had a less than 70 percent LPA audience in either 2004 or 2005) and it had alcohol advertising prior to the 2003 rule change. Eight television programs met these criteria: *106 & Park*, *Behind the Music*, *BET Comic View*, *Crank Yankers*, *Crocodile Hunter*, *Driven*, *Rap City*, and *BET Top 25 Countdown*. Five magazines also met the criteria: *Allure*, *Automobile*, *ESPN*, *Spin*, and *Vibe*.

Relatively few programs and magazines met both criteria. Most programs and magazines with LPA audiences of less than 70 percent do not carry alcohol advertising. These media either so obviously appeal to youth (e.g., *Spongebob Squarepants* and *Seventeen Magazine*) or appear to appeal to youth (e.g., adult oriented cartoons such as *The Simpsons*) that alcohol firms avoid advertising in them or the media refuse alcohol advertisements. Furthermore, many media popular with alcohol advertisers, such as sports programming, have such broad

[10] Mediamark Research & Intelligence is the other source.

[11] I only have access to television ratings data for 2000-2005. I assume programs "ineligible" for alcohol advertising in 2005 were also "ineligible" in 2006.

appeal across all ages that they are well above the 70 percent threshold. Given the alcohol industry's acquiescence to the FTC's recommendations, perhaps it is unsurprising that the rule change imposed little cost in terms of advertising freedom.

Figure 1.4 plots the total number of advertisements that appeared each year on programs and in magazines affected by the 2003 rule change. Advertising in these magazines and programs drops off dramatically after 2003. However, advertising declines more clearly between 2003 and 2004 in magazines than on television. Additionally, many of these programs and magazines continue to carry advertisements despite the rule change, though all show fewer ads. For each selected program and magazine, I create an indicator variable with a value of one if the magazine or program was ineligible for alcohol advertising in a given year (i.e., above 70 percent LPA audience after 2003) and zero if the magazine or program was eligible for alcohol advertising (i.e., above or below 70 percent prior to 2003 or above 70 percent after 2003). I isolate variation associated with the rule by using these indicators as instruments for exposure to alcohol advertising, as described in more detail below.

Measures of Individual Alcohol Advertising Exposure

The richness of the NCS media consumption data and the Kantar Media advertising data allows me to create comprehensive and individual estimates of exposure to both magazine and television alcohol advertisements. Following Avery et al. (2007), I generate estimates of alcohol magazine advertising exposure using data on each individual's magazine reading practices. Each NCS respondent indicates the fraction of the last four issues he read of each of 173 magazines. I assume this fraction proxies for the fraction of issues of a magazine the respondent read over the past six months. I multiply the fraction of magazine issues a respondent reads by the number of alcohol advertisements appearing in that magazine in the six months prior to the start of the survey[12] to estimate the number of advertisements he saw in that magazine. I then sum these estimates over all magazines he reports reading to create his total potential exposure to magazine alcohol advertising. Formally, potential exposure to alcohol advertisements of respondent i in year t is given by:

$$ad_{it}^{mag} = \sum_{m} view_{imt} \times ads_{mt} \tag{1.1}$$

where $view_{imt}$ denotes the reading intensity variable, and the subscripts denote respondent i, magazine m, and wave t.[13] Similar to Avery et al. (2007), I assume that advertising does not depreciate over time until six months after exposure when it depreciates completely. While this is a simplifying assumption, a review of the literature studying advertising depreciation by Leone (1995, pg. 149) finds "evidence that the average expected carryover effect is between six and nine months."

Similar to the magazine advertising exposure measure, the television measure uses survey information on which programs an individual reports viewing. Respondents report how often they watch a long list of regularly occurring broadcast, cable, and sports programs. Parallel to (1), I generate a probability that a respondent viewed a specific airing of each program and weight the number of ads appearing on the program in the past six months by this probability of viewing. The NCS also asks respondents if they viewed one-time special programming (e.g., the Academy Awards, the World Series). For each special program a person says he watched I assume he saw all alcohol advertisements that aired during that show. I then sum over all of the regular programs and specials he reports watching to create his total potential exposure to television advertising in the six months prior to the survey. Appendix B explains

[12] Simmons surveys NCS respondents over a four month window for each wave. The NCS data do not indicate when in this period an individual answers and returns the survey, introducing some measurement error due to the imprecise matching of advertising exposure and drinking.

[13] Note the individual data are cross-sectional.

the advertising exposure estimates in more detail.

1.3.1 Summary Statistics

Table 1.6 presents summary statistics for the main variables in the models. I find that on average 36 percent of under-age youth, ages 18 to 20, and 64 percent of over-age youth, ages 21 to 24, drank any alcohol in the last 30 days. Under-age drinkers consumed 22 drinks in the last 30 days and 22 percent drank “heavily” (more than 33 drinks). Over-age drinkers drank 24 drinks and 23 percent drank “heavily.” These consumption trends are similar, but slightly lower than, those from the National Survey on Drug Use and Health (NSDUH). Appendix A further compares alcohol consumption trends in the NCS and NSDUH. The average 18 to 24 year old saw 725 television alcohol advertisements and 70 magazine alcohol advertisements over a six month period. Though these exposure averages are larger than previous measures, particularly for television advertising, they match temporal trends in advertising expenditures and other alcohol advertising exposure measure.[14] Those who drank saw 315 more television and 24 more magazine advertisements than those abstaining from alcohol. On average a respondent saw 5 advertisements in magazines the 70 percent LPA audience rule affected and 18 ads on affected television programs.

1.4 Methods and Identification

Alcohol firms’ targeting of advertisements at those who already drink greatly complicates establishing whether alcohol advertisements cause youth to drink. Though numerous studies have demonstrated a strong positive relationship between alcohol advertising and youth drinking, they largely ignore targeting. The primary contribution of this study is a serious

[14] Appendix B compares the advertising exposure measures I use in this paper with other measures of advertising and advertising exposure.

attempt to confront targeting's threat to establishing a causal link between advertising and drinking.

An alcohol firm has three ways to increase sales by advertising. First, a firm can convince drinkers of other brands to try its brand. Second, the firm can convince drinkers of its brand to continue drinking the brand or drink more of it. Finally, the firm can convince non-drinkers to not only begin drinking but also to drink its brand. Each of these potential goals suggests a different advertising strategy. To accomplish the first goal, firms should target advertisements at consumers of other brands. The second goal implies firms should target advertisements at their own consumers. The third implies targeting non-drinkers. However, these strategies are certainly not mutually exclusive and a profit maximizing firm should employ each to the point where the expected marginal profit from the last persuaded consumer is equal to the marginal advertising cost. The first two strategies imply that drinkers see more advertisements than non-drinkers, as firms seek media popular with drinkers to fight for market share. The third strategy implies non-drinkers see more advertisements. Not properly controlling for targeting will bias estimates of the effect of advertising on drinking. If the targeting of drinkers prevails, the bias will be positive (i.e., toward finding an advertising effect), and if the targeting of non-drinkers dominates, the bias will be negative.

As an illustrative example of the potential bias due to targeting, consider two primetime, broadcast television programs: *NYPD Blue* and *Touched by an Angel* (See Table 1.6). According to the NCS in 2003, 70 percent of *NYPD Blue* viewers drank in the last 30 days, and, conditional on drinking, consumed an average 16 drinks in that period. Only 46 percent of viewers of *Touched by an Angel* drank, and the average drinker consumed 10 drinks. In 2003, 42 alcohol advertisements aired on *NYPD Blue*, but *none* appeared on *Touched by an Angel*. Table 1.6 also compares readers of *Reader's Digest* and *Rolling Stone*. Rolling Stone readers are much more likely to drink than *Reader's Digest* readers, and Rolling Stone contains 148 alcohol advertisements in 2003 compared to 1 in *Reader's Digest*. Though two programs and two magazines are a small sample, this general relationship between advertising and the drinking of viewers holds for other programs and magazines. It seems unlikely that this large disparity in drinking behavior is driven entirely by an effect of advertising on behavior, rather than an attempt by alcohol firms to place advertising where drinkers watch. Additionally, though some of these differences in advertising levels may be driven by demographic differences in these audiences, controlling for these differences may only exacerbate the selection bias. A 50-year-old *NYPD Blue* viewer may not drink much more than a 50-year-old *Touched by an Angel* viewer, but there may be much larger differences between 18-year-old viewers of these programs. This example illustrates that the endogeneity of advertising exposure, driven by media choice and targeting, is a real and significant concern. This paper uses several methods to control for this endogeneity.

Model Specification

I estimate three models of drinking behavior: (1) the probability a person drank any alcohol in the 30 days prior to the survey, (2) the number of drinks she consumed in the past 30 days conditional on drinking, and (3) the probability a person drank "heavily"[15] conditional on drinking. The main covariate of interest measures the relationship between alcohol advertising exposure and these behaviors. I estimate separate effects of magazine and television advertising to allow for different effects of advertising in each medium. I also interact each of these measures with an under-age indicator variable to allow the effect to differ for under-age

[15] I define "heavy" drinking as consuming more than 33 drinks in the last 30 days. I base this definition on the average number of drinking days for this age group (6.7) times the definition of binge drinking (5+ drinks on one occasion). The definition roughly corresponds to the 75th percentile of consumption for youth drinkers in the NCS.

and over-age youth. The general structure of the the model is:

$$
\begin{aligned}
drk_{it} = \beta_0 + \beta_1 ad_{it}^{tv} + \beta_2 ad_{it}^{tv} uage_{it} + \beta_3 ad_{it}^{mag} + \beta_4 ad_{it}^{mag} uage_{it} \\
+ demo_{it}\beta_5 + media_{it}\beta_6 + T_t\beta_7 + M_m\beta_8 + \epsilon_{it}
\end{aligned}
\tag{1.2}
$$

where i indexes an individual, t refers to a survey wave[16], and m is a media market. drk_{it} is either a 0 or 1 indicator of any past 30 day drinking, a continuous measure of the number of drinks in the last 30 days, or an indicator of "heavy" drinking. ad_{it}^{tv} and ad_{it}^{mag} are measures of potential television and magazine advertising exposure. $demo_{it}$ is a vector of standard demographic characteristics: age, race, gender, household income, employment and student status, education, religion, household characteristics, and the number of adults in the household. $media_{it}$ is a vector of controls for media choice and targeting that vary by specification. I first estimate a simple specification with few controls for media choice, then I sequentially add stricter controls. T_t is a set of wave fixed effects that control for changes in average drinking over time, and M_m is a set of media market fixed effects that control for differences in average drinking across DMAs.

My simple specification is similar to previous studies, as it includes only basic controls of media consumption: the number of magazine issues the individual read in the last six months and the number of hours spent watching television in an average week. However, unlike most previous studies, even this simple specification includes both market and time fixed effects. I estimate this naive, simple model to demonstrate the potential bias due to omitting controls for targeting. This specification uses three sources of variation in advertising exposure: program and magazine viewing choices, advertising placement in those magazines and programs over time, and differences in placement in the same period across media markets (television only). Because I control for general across-time and across-market differences, the variation

[16] The NCS is a repeated cross-section and I observe each individual only once. I include the time subscript to denote temporal variation in advertising.

in my simple model is driven by the within-market differences in advertising over time and by individual media choices. To interpret these model results as casual, one must assume that firms do not target DMAs over time and, more importantly, that firms do not target advertising at certain programs and magazines.

Next, I extend the simple model to include very basic controls meant to proxy for the likelihood that alcohol firms target the programs and magazines one views. Firms cannot target television and magazine advertising at specific individuals.[17] They can only plausibly target advertising at particular magazines and television programs. Therefore, I attempt to control for the probability alcohol advertisers target the media someone views by generating summary statistics of the audiences of the media she watches. Essentially, these models control for whether, on average, one views media popular with drinkers, young people, and other groups. To generate these controls, I use the NCS to create average reader and viewer demographic characteristics for each magazine and television program: race (percent Black, Asian, Hispanic, and other race), education (percent high school dropout, some college, and bachelors degree), household size, real household income, age (average age and percent underage and age 21 to 24), and sex (percent female). I also create the average alcohol consumption for each program and magazine, including: percent of readers/viewers who drank in the last 30 days, the number of drinks in the last 30 days for drinkers, and the number of those drinks that were beer and liquor. For each program and magazine a respondent views, I match him to the medium's audience characteristics from the previous NCS wave to avoid simultaneity and account for firms using prior information to target in the current period. I then average these audience characteristics over all magazines he reads to create the average audience of the magazines he reads. Similarly, I average over all television program he views to create his average television audience characteristics. Intuitively, these controls measure

[17] Firms can more finely target internet advertising. However, in the period I study internet advertising is a very small fraction of alcohol advertising expenditures (Adams Beverage Group, 2006*a*,*b*,*c*).

the probability one drinks and how much they likely drink based solely on the media they view.

If firms use these demographic and drinking measures to target advertising, these measure control for the likelihood firms target the average magazine or television show one views. However, this approach assumes these media viewing averages are sufficient statistics for the targeting of an individual. If these averages do not completely capture firms' targeting, omitted variable bias will remain. Despite this limitation, these simple targeting controls more strictly control for endogeneity than any previous study of alcohol advertising.

An approach to addressing targeting that requires fewer assumptions is to control directly for media choice. In these models, I include television program and magazine choice controls that measure the frequency with which a person reads each magazine and watches each program listed in the NCS. These measures are exactly what I use to create exposure estimates (see Equation 1). The choice controls account for any variation in the exposure to alcohol advertising due to individual media choice, leaving only variation occurring within a program or magazine over time or within a program across markets.[18] To interpret the results of these models as causal, one must assume that alcohol firms do not continuously adjust the number of advertisements on a program or in a magazine as they receive new information about its viewers. While this is a weaker assumption than believing firms do not target programs at all, it also may be invalid. Additionally, it is a very demanding econometric specification, requiring estimating 2,510 television program choice controls and 173 magazine choice controls. The remaining variation in advertising exposure may not be sufficient to precisely estimate the effect of advertising. However, this specification places few assumptions on the targeting behavior of firms and uses decidedly stronger controls for endogeneity than in previous work.

[18]My exposure measures only include within wave variation across media markets for broadcast television programs.

LPA Audience Instruments

As a final and more internally valid method to control for targeting, I isolate variation in advertising due to the 2003 change in the alcohol industry's legal purchasing age (LPA) audience rule. This variation is exogenous under plausible assumptions and allows me to address the possibility that firms may target on changing, unobserved program characteristics over time. Intuitively, I compare two people who watch the same program, but one sees many fewer advertisements on that program after 2003 due to the stricter rule. This approach controls for targeting by using a known and exogenous source of variation in the amount of advertising people see.

I isolate variation in advertising due to the rule change by using the viewing of media affected by the rule change as instruments for advertising exposure. As described above (Section 3), the change to a stricter 70 percent LPA audience rule directly impacted 8 television programs and 5 magazines. For each of these 13 media, I create a variable indicating whether the rule applied to that program in a given year. That is – the rule applied to a program in 2004-2007 and the program had an audience of less than 70 percent LPA in the prior year. I then interact this indicator with a variable that measures how frequently one views that magazine or program. This interaction variable equals the frequency one views a program or magazine in years the 70 percent rule was binding for that program or magazine (2004-2007). The interaction equals zero in years before the 2003 rule change (2001-2003), in years after the rule change if the program or magazine had an above 70 percent LPA audience, and for respondents who did not watch or read the program or magazine. Formally, the instrument for each medium is:

$$Ins_{it}^{k} = 1\{t > 2003\} \times 1\{LPA_{t-1}^{k} < 70\%\} \times view_{it}^{k} \tag{1.3}$$

where LPA^k_{t-1} is the percentage of the audience of medium k of legal purchase age (LPA) in year *t-1* and $view_{imt}$ denotes fraction of issues/airings of that medium respondent i viewed in year t. I use these instruments in the first stage of a two-stage least squares model to predict alcohol advertising exposure. To estimate a differential effect for youth, I also create a second set of instruments by interacting the first set of instruments described above with a variable indicating a respondent is under-age.[19] I use both sets of instruments in four first-stage estimations predicting: magazine advertising exposure, magazine exposure for the under-age, television exposure, and television exposure for the under-age. To control for the endogeneity of watching these 13 media, I also include variables measuring the frequency of viewing/reading for each medium in both the first and second stages. The exclusion restriction assumption is that viewers of these programs and magazines have the same underlying propensity to drink in years when the rule bars ads as they do in years the rule does not bar ads. The strength of these programs and magazines as instruments rests on whether the LPA audience rule is binding for advertisers.

To properly identify the effect of these rules on these specific magazines and programs, I split my advertising exposure measures into two separate measures: exposure in magazines and programs affected by the rule change and exposure in media unaffected directly by the rule change.[20] Splitting the exposure measures is necessary because viewing programs and magazines affected by the rule change is correlated with viewing media unaffected by the change. Advertising in media unaffected by the rule increased after 2003, due to both a general upward trend in alcohol advertising and the possible general equilibrium effect of removing some alcohol advertising supply (the affected media) from the market. Without controlling for exposure in media the rule does not directly affect, the instruments may

[19]This interaction avoids estimating the "forbidden regression"(Angrist and Pischke, 2008, pg. 190-192). That is, one estimates the second stage using the interaction of predicted exposure and an under-age indicator.

[20]I outline the method I use to create exposure estimates in the Data section, and I discuss it in further detail in Appendix B.

predict that those who watch affected programs after the rule change actually see more ads, rather than fewer.

As a brief example, viewing of *Crank Yankers* may be highly correlated with viewing of the *The Daily Show with Jon Stewart*. Both shows appear on *Comedy Central* and typically aired consecutively. While the 2003 rule change made *Crank Yankers* ineligible for alcohol ads, it did not affect *The Daily Show*, which experienced large increases in alcohol advertising in 2004 and 2005. If enough people view both programs and the advertising increase on *The Daily Show* is large enough, then the average *Crank Yankers* viewer will see more ads after the rule change rather than less. However, they will still see fewer ads specifically on *Crank Yankers*. To properly isolate the effect of the rule change on *Crank Yankers* viewers, I need to control for exposure on other programs, including *The Daily Show*.

Controlling for exposure in media unaffected directly retains the intuition described earlier; I compare two people who saw the same number of ads in all other media but saw very different amounts in the 13 affected programs and magazines. I assume the difference in advertising in the 13 media is driven by the LPA audience rule change. Formally, I estimate the general first-stage equations:

$$insad_{it}^{tv} = \pi_0 + \pi_1 Ins_{it} + \pi_2 InsFE_{it} + \pi_3 nonad_{it} + demo_{it}\pi_4 + T_t\pi_5 + M_m\pi_6 + \zeta_{it} \quad (1.4)$$

$$insad_{it}^{mag} = \phi_0 + \phi_1 Ins_{it} + \phi_2 InsFE_{it} + \phi_3 nonad_{it} + demo_{it}\phi_4 + T_t\phi_5 + M_m\phi_6 + \zeta_{it}, \quad (1.5)$$

where $insad^{tv}$ and $insad^{mag}$ are, respectively, the estimated exposures to advertising on programs or magazines unaffected directly by the 2003 rule. Ins_{it} is the vector of instruments measuring whether an individual viewed an affected medium when it was ineligible for advertising; it is excluded from the second-stage equation. $InsFE_{it}$ is a vector of fixed effects measuring the general viewing of affected programs; it remains in the second stage to control for the underlying propensity to drink for viewers of the 13 affected programs and magazines. $nonad_{it}$ is a vector of measures of exposure to advertisements in unaffected media and is included in both the first- and second-stages. I estimate a model using two-stage least squares. The second stage is similar to equation (2) and includes both the predicted exposure to advertising on affected media and the estimated exposure in unaffected media. However, I only attach a causal interpretation to the predicted exposure generated by my instruments.

The estimates of this instrumental variables model represent the local average treatment effect (LATE) of advertisements on these 8 programs and in these 5 magazines on their viewers and readers. These estimates are more generally informative under two conditions: the advertisements on these programs and magazines are similar to ads in other programs and magazines and the viewers and readers of these programs and magazines are similar to other youth. The advertisements in these programs and magazine are similar, if not identical, to advertising appearing on other programs. However, it is difficult, if not impossible, to test whether advertising and programming may complement one another, making the same ad more effective on one program than another.[21] Though the rule only affected 13 programs and magazines, these media are very popular with youth. Almost half (49.9 percent) of my sample viewed at least one of these programs or magazines. I compare viewers and non-viewers of these 13 media in Appendix C. I also compare these viewers before and after the 2003 LPA rule change. I find that with some small exceptions the viewers of these 13 programs and magazines look like non-viewers and do not differ substantively after the rule change. These similarities imply that the local average treatment effect of advertising on these 8 programs and 5 magazines is similar to the effect of alcohol advertising on other programs and in other magazines. Additionally, because the LATE of this model estimates the effect on youth drinking of a policy meant to reduce youth alcohol advertising exposure,

[21] This assumes the audiences of these two programs are the same.

the LATE estimate itself is interesting.

1.5 Results

Table 1.6 reports selected coefficients from models of past 30 day drinking prevalence. The first column reports results for the baseline model without any controls for targeting. The main magazine and television advertising coefficients give the relationship between advertising in that medium and drinking for over-age youth, and the sum of each main coefficient and the coefficient on the interaction term gives the relationship for under-age youth. I conduct tests of joint significance (F-tests) for each under-age relationship and indicate the level of joint significance in my tables. The estimates demonstrate a strong relationship between alcohol advertising exposure and drinking. The coefficients imply that seeing 100 additional magazine ads is associated with an increase in the probability an under-age youth drinks of 6.2 percentage points and an increase in the probability an over-age youth drinks of 6.6 percentage points. An additional 100 television ads is related to a 0.2 percentage point increase in the probability of drinking for the under-age and a 0.3 percentage point increase for the over-age. If interpreted as causal, these results imply a fifty percent decrease in the average magazine and television exposure would result in a 3.2 percentage point decrease in the under-age drinking prevalence and a 3.5 percentage point decrease in over-age drinking. These results are similar to those found in Saffer and Dave (2006).[22]

The model in the second column of Table 1.6 adds controls for the average demographics and drinking behavior of the programs and magazines a respondent views. While the coefficient estimates are similar to the baseline model, they are smaller, which is consistent with findings that readers of magazines that are more popular with drinkers are more likely

[22] Saffer and Dave (2006) report a 28 percent reduction in ad exposure is associated with a 1 to 4 percentage point drop in the drinking prevalence of 14 to 18 year olds. The results in column 1 of Table 1.6 indicate the same reduction in ad exposure implies a 2.8 percentage point decrease in 18 to 20 year old drinking prevalence.

to drink themselves. The controls measuring the probability one drinks given the magazines and television programs he reads and watches are highly predictive of one's own drinking, implying media choice is substantially related to drinking. These results provide a conservative upper-bound estimate of the effect of advertising on youth drinking. While these estimates control for targeting better than the previous literature, they still may contain some bias due to targeting. The media audience average results imply that a fifty percent reduction in the average youth alcohol advertising exposure would decrease under-age drinking prevalence from 36.3 percent to 34.3 percent and decrease over-age drinking prevalence from 64.0 percent to 61.7 percent. These effects are notably smaller than those from column 1, particularly for under-age youth.

Column 3 of Table 1.6 reports selected coefficient estimates from models using magazine and television program choice controls. After including strong controls for media choice, the effect of television and magazine advertising on under-age and over-age drinking is no longer statistically significant. Although I cannot rule out larger effects, the point estimates imply that a fifty percent decrease in the average magazine and television advertising exposure would *increase* under-age drinking prevalence by 0.6 percentage points and over-age drinking prevalence by 0.2 percentage points. These effects are much smaller than those from columns 1 and 2 and imply that a large reduction in youth alcohol advertising exposure will have negligible effects on youth drinking prevalence.

The fourth column of Table 1.6 reports results from the two-stage least squares model. I find that the television program advertising eligibility instruments are strong predictors of advertising exposure in those media (See Appendix C for first-stage results) with combined F-statistics of at least 48(Stock and Yogo, 2001). However, the magazine advertising eligibility instruments are relatively weak predictors of exposure in affected magazines, which may lead to imprecise second stage estimates of the effect of magazine advertising. The second stage results show a positive general effect of both magazine and television advertising and a larger negative added effect for youth. While imprecisely estimated, these results imply a negative effect of magazine advertising on the over-age and a net positive effect on the under-age. The television results are the opposite of the magazine results with a net positive effect for the over-age and a net negative effect for the under-age. However, these results are not statistically different from zero. The coefficients estimating the relationship between non-instrumented media and drinking are still highly significant and similar to the results in column 1.

Table 1.6 is analogous to Table 1.6 and presents the results of models of past 30 day drinking volume conditional on consuming at least one drink. These results show a similar pattern to Table 1.6. The coefficients in columns 1 and 2 show a strong relationship between drinking volume and advertising exposure. The results in column 1 imply seeing an additional 100 magazine ads is associated with an increase of 4.0 drinks per month among under-age drinkers and 4.4 drinks among over-age drinkers. An additional 100 television ads is associated with an increase of 0.39 drinks per month for under-age drinkers and 0.41 drinks for over-age drinkers. The magazine and program choice control estimates in column 3 are distinct from those in columns 1 and 2. The results imply a fifty percent decrease in average magazine and television advertising exposure would results in a decrease of an average one drink per month for the under-age and 1.2 drinks for the over-age. However, the standard errors of these coefficients are quite larger than in columns 1 and 2, and none of the coefficients are statistically significant. The two-stage least squares results in column 4 also show a similar pattern to Table 1.6 (See Appendix C for first-stage results). The coefficient estimates are much larger than in the other three models of consumption volume, particularly for television exposure. The results imply a large, positive effect of advertising for over-age drinkers and an even larger, negative effect for under-age drinkers. Although the confidence intervals include large effects, none of these coefficients are statistically significant.

Table 1.6 displays results of models of "heavy" past 30 day drinking, conditional on drinking. The coefficient estimates in column 1 imply that seeing an additional 100 magazine advertisements is associated with a 5.3 percentage point increase in the probability an under-age drinker drinks "heavily" and a 6.2 percentage point increase for over-age drinkers. Viewing an additional 100 television advertisements is associated with a 0.4 percentage point increase in "heavy" drinking among under-age drinkers and a 0.5 percentage point increase among over-age drinkers. Though noisy, the media choice controls model estimates in column 3 imply a fifty percent reduction in average alcohol advertising exposure would *increase* under-age "heavy" drinking by 0.3 percentage points and decrease over-age "heavy" drinking by 0.1 percentage points. The two-stage least squares results in column 4, similar to those in tables 3 and 4, are typically much larger than those in columns 1 to 3. The results in column 4 imply large, perhaps implausible, effects of magazine advertising and similarly large effects of television advertising on under-age and over-age "heavy" drinking.

1.6 Discussion

Though numerous studies establish a strong correlation between advertising and drinking, credibly demonstrating a casual effect is notably more difficult. A person does not randomly choose television shows or magazines. These choices likely reveal more about a person's interests than would knowledge of his age, race, education, or income. Advertisers understand programs and magazines provide a diversity of audiences and probably target advertising at audiences with the most interest in their product. I use data and methods particularly well suited to accounting for targeting and identifying any remaining effect of advertising on behavior. I use summary statistics of the media one consumes, allowing for simple targeting behavior by firms. I also use an agnostic, "brute force" method, controlling for each magazine and program a person views. Finally, I use a more fine-tuned approach, exploiting a source of variation in advertising placement that is plausibly not driven by targeting. The resulting evidence confirms a strong superficial relationship between alcohol advertising and youth drinking, but it does not support an interpretation that the relationship is causal. This strong correlation is not robust to the two strictest methods I use to control for targeting.

Even though I find a strong unconditional relationship between advertising and drinking, more credible estimates of this relationship imply a small or null economic impact. Though the media choice controls and instrumental variables models use the strongest controls for targeting, estimates from these models are imprecise. The media audience average estimates (column 2) provide a conservative upper-bound estimate of the effect of advertising on youth drinking. These estimates imply smaller advertising effects than those found in much of the literature. Although these estimates are more precisely estimated than estimates from models with stricter targeting controls, they may be biased by targeting. Estimates from more credible models with media choice controls imply that such a reduction in advertising exposure will have negligible effects on all three measures of youth drinking.

Evidence from the media choice controls and instrumental variables models, specifications that are statistically more credible, is notably less precise. This result is unsurprising since these models use much less residual variation in advertising exposure to estimate its effect on drinking. The confidence intervals of these results do not rule out large effects of alcohol advertising. However, these large effects seem unlikely given the results of the less sophisticated models. Such a difference would imply a large negative targeting bias (i.e. firms targeting media with relatively fewer drinkers). Furthermore, though the estimates of the media choice controls and instrumental variables models are noisy, the purpose of estimating them is to determine whether the data support rejecting the null hypothesis that advertising has an effect in a model based on credible assumptions. These two models require fewer and more credible assumptions about the exogeneity of advertising exposure than anything in the literature. However, their results show no general evidence to support a causal effect of alcohol advertising on youth consumption.

An interesting and consistent result is the smaller advertising relationship for under-age youth compared to those over-age. This finding might stem from firms not targeting advertising at the under-age, or at least targeting them less. Similarly, among the under-age there may be less selection into different media based on underlying characteristics. The under-age may watch generally the same group of shows regardless of whether they drink, leading to a weaker relationship between advertising and drinking. Another possibility is that the compositions of under-age and over-age drinkers differ from each other. The 68 percent of 21 to 24 year olds who drink is composed of those who drank while under-age and those who did not. These law-abiding drinkers, who waited until age 21 to drink, may be less committed to either drinking or abstaining and perhaps more susceptible to advertising on the margin. Under-age youth may simply be less affected by alcohol advertisements. According to industry advertising codes, advertisements must avoid appealing to the under-age. Perhaps the industry has been successful in appealing more to the over-age. Finally, under-age youth may face constraints on acting on advertising, because they cannot legally purchase alcohol. Regardless of why, if alcohol advertising affects the decisions of whether and how much to drink, my results suggest that it is more effective on over-age youth than on those under-age.

Although this study improves the estimates of the effect of alcohol advertising on youth drinking, it has several limitations. The NCS does not allow me to study "problem" drinking behaviors such as binge drinking or drunk driving. The illegality of under-age drinking elicits concern about a link between advertising and any level of under-age consumption. However, judging a possible link between advertising and legal drinking is more problematic. Some "heavy" drinkers in this study could be consuming healthy, perhaps even beneficial, amounts of alcohol if they have a drink or two every day. Without establishing a link between advertising and "problematic" over-age drinking this study cannot claim that this relationship, if present, is a public health concern. Additionally, this study does not address potential longer term effects of alcohol advertising exposure. If alcohol advertising acts on a longer time line, perhaps starting in childhood, these data cannot identify this effect.

Overall, this study suggests that alcohol advertising is, at most, a small influence on under-age drinking. Advertising may play a larger role in the drinking behavior of over-age youth. However, more credible estimates find no support for a casual effect of advertising on youth drinking. National trends in alcohol advertising spending and youth drinking support this small to null advertising effect. Though several studies, including this one, find large increases in alcohol advertising spending from 2001-2005,[23] under-age drinking prevalence remains nearly constant over this period (See NSDUH in Figure 1.6). While it is possible that drinking prevalence would have declined if not for advertising, these trends do not provide evidence of a strong effect of advertising. Though the lack of evidence of a causal effect of alcohol advertising may disappoint those looking for ways to reduce the problems associated with under-age drinking, the glass may be half full. It appears that at the least industry self-regulation is not failing, although this study does not establish if industry self-regulation is the reason youth do not react to advertising. Furthermore, stronger restrictions on alcohol advertising, voluntary or legal, might be difficult to establish due to unpopularity or constitutional concerns. The results of this study suggest such policies curtailing or eliminating alcohol advertising would have little or no effect on under-age drinking.

[23]Spending on television and magazine alcohol advertising increased 31.1 percent and 17.9 percent from 2001 to 2005 respectively (Kantar Media).

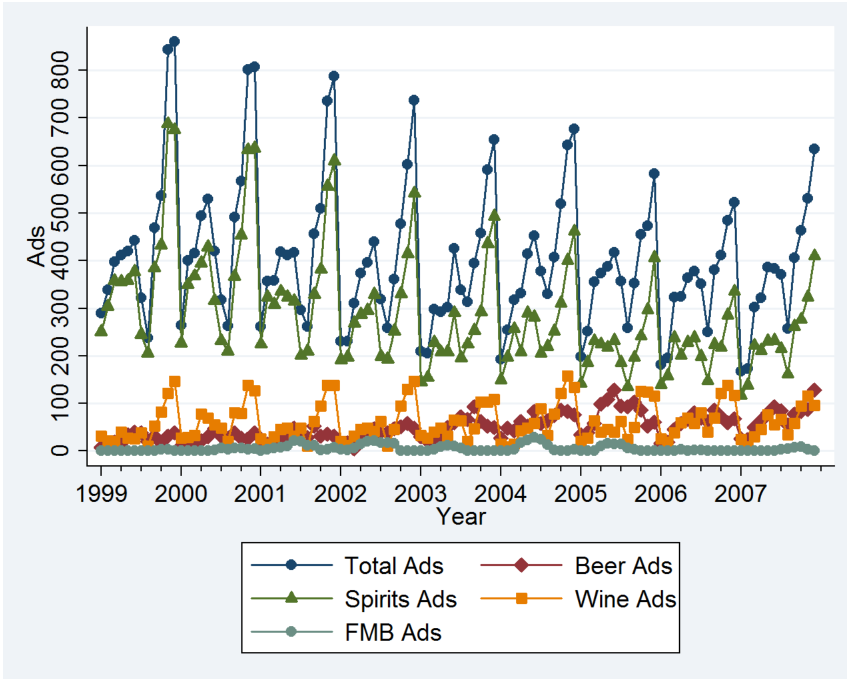

Figure 1.1: Number of alcohol advertisements, both in total and by alcohol type, in all magazines tracked by Kantar Media, 1999-2007. Note: FMB abbreviates flavored malt beverages (e.g., *Smirnoff Ice* and *Mike's Hard Lemonade*). Source: Kantar Media.

Table 1.1: Comparison of Selected TV Program Viewers and Advertising in 2003

Program/Magazine	Pct. Drink	Avg. Drinks/30 Days	Alcohol Ads
NYPD Blue	70	16	42
Touched by an Angel	46	10	0
Rolling Stone	72	20	148
Reader's Digest	59	11	1

Source: 2003 NCS and Kantar Media.

Table 1.2: Summary Statistics

	Full Sample Mean	SD	Drinkers Mean	SD
Any Drinks in L30 Days	0.52		1.00	
Drinks in L30 Days	11.90	21.23	23.10	24.82
Heavy Drinking	0.12		0.23	
Mag Ad Exposure	0.70	1.04	0.84	1.09
Mag Ad Exposure (Instrumented Media)	0.05	0.12	0.05	0.13
TV Ad Exposure	7.25	11.05	8.24	11.39
TV Ad Exposure (Instrumented Media)	0.18	0.47	0.22	0.50
Female	0.55		0.53	
Educ: Less than HS	0.19		0.17	
Educ: Some College	0.34		0.36	
Educ: College	0.13		0.17	
Full-Time Student	0.26		0.23	
Part-Time Student	0.07		0.08	
Employed Full-Time	0.42		0.50	
Employed Part-Time	0.25		0.24	
Married	0.39		0.37	
Age	20.85	1.99	21.39	1.84
Under-age	0.45		0.32	
Hispanic	0.42		0.39	
Black	0.08		0.07	
Other Race	0.17		0.15	
Number of Adults in HH	3.36	1.23	3.28	1.21
Real HH Income (000)	86.82	66.16	90.07	67.95
Likely Parent in HH	0.74		0.72	
Spanish Predom Lang in HH	0.25		0.23	
Other Lang Predom in HH	0.03		0.03	
Mag Issues Read	99.65	120.58	110.56	120.11
Total Count of Netwk Shows	15.96	26.21	17.08	25.81
Hours spent viewing netwk TV per week, sum of daypts	15.75	16.21	16.46	16.20
Total Hours viewing cable TV per week	15.07	23.90	16.53	24.81
Total Count of cable Shows watched	16.54	30.26	18.04	30.26
Read No Magazines	0.13		0.10	
Watch No TV	0.04		0.03	
Observations	12752		6568	

Table 1.3: LPM/2SLS Estimated Effect of Advertising on 30 Day Drinking

	(1) Baseline	(2) Media Avg Aud	(3) Media FE	(4) IV
Mag Ad Exposure	0.0657***	0.0349***	-0.0272	
	(0.0083)	(0.0094)	(0.0264)	
× Under-age	-0.0036	-0.0017	-0.0009	
	(0.0083)	(0.0083)	(0.0111)	
TV Ad Exposure	0.0033***	0.0028***	0.0022	
	(0.0008)	(0.0008)	(0.0025)	
× Under-age	-0.0013*	-0.0013	-0.0025**	
	(0.0008)	(0.0008)	(0.0012)	
Mag Ad Exposure (Instrumented Media)				-0.0750
				(0.4626)
× Under-age				-0.2496
				(0.8319)
TV Ad Exposure (Instrumented Media)				0.0779
				(0.0508)
× Under-age				-0.0412
				(0.0766)
Mag Ad Exposure (Non-Ins Media)				0.0731***
				(0.0090)
× Under-age				-0.0142
				(0.0109)
TV Ad Exposure (Non-Ins Media)				0.0036***
				(0.0008)
× Under-age				-0.0014
				(0.0010)
Observations	12752	12752	12752	12752
R^2	0.169	0.184	0.415	0.173
F-Statistic (Mag Ad Exposure)				8.88
F-Statistic (× Under-age)				5.72
F-Statistic (TV Ad Exposure)				82.04
F-Statistic (× Under-age)				48.40

Robust standard errors in parentheses; All advertising measures divided by 100
All models include demographic controls, magazine issues read in the past 6 months, hours of TV viewed per week, and market (DMA) and survey wave (time) fixed effects.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Joint statitistical significance of under-age effect: $^{+}$ $p < 0.10$, $^{++}$ $p < 0.05$, $^{+++}$ $p < 0.01$

Table 1.4: OLS/2SLS Estimated Effect of Advertising on 30 Day Consumption Volume

	(1) Baseline	(2) Media Avg Aud	(3) Media FE	(4) IV
Mag Ad Exposure	4.4429***	3.7110***	2.8379	
	(0.5272)	(0.6114)	(2.0456)	
× Under-age	-0.4335	-0.5242	-0.9958	
	(0.5916)	(0.5912)	(0.9285)	
TV Ad Exposure	0.4135***	0.3901***	0.0047	
	(0.0509)	(0.0533)	(0.2052)	
× Under-age	-0.0254	0.0024	-0.0246	
	(0.0555)	(0.0555)	(0.1075)	
Mag Ad Exposure (Instrumented Media)				-7.0360
				(36.1684)
× Under-age				-33.6649
				(51.0261)
TV Ad Exposure (Instrumented Media)				2.0500
				(3.8025)
× Under-age				-4.9699
				(5.9761)
Mag Ad Exposure (Non-Ins Media)				4.3176***
				(0.7236)
× Under-age				-0.1350
				(0.8458)
TV Ad Exposure (Non-Ins Media)				0.4145***
				(0.0755)
× Under-age				-0.0761
				(0.0960)
Observations	6568	6568	6568	6568
R^2	0.123	0.138	0.679	0.131
F-Statistic (Mag Ad Exposure)				7.48
F-Statistic (× Under-age)				4.09
F-Statistic (TV Ad Exposure)				50.67
F-Statistic (× Under-age)				25.13

Robust standard errors in parentheses; All advertising measures divided by 100
All models include demographic controls, magazine issues read in the past 6 months,
hours of TV viewed per week, and market (DMA) and survey wave (time) fixed effects.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Joint statitistical significance of under-age effect: $^{+}$ $p < 0.10$, $^{++}$ $p < 0.05$, $^{+++}$ $p < 0.01$

Table 1.5: LPM/2SLS Estimated Effect of Advertising on Drinking

	(1) Baseline	(2) Media Avg Aud	(3) Media FE	(4) IV
Mag Ad Exposure	0.0662***	0.0516***	0.0178	
	(0.0091)	(0.0106)	(0.0378)	
× Under-age	-0.0131	-0.0140	-0.0183	
	(0.0102)	(0.0102)	(0.0171)	
TV Ad Exposure	0.0051***	0.0046***	-0.0015	
	(0.0009)	(0.0009)	(0.0038)	
× Under-age	-0.0009	-0.0005	-0.0011	
	(0.0010)	(0.0010)	(0.0020)	
Mag Ad Exposure (Instrumented Media)				0.4989
				(0.5786)
× Under-age				-1.0372
				(0.7876)
TV Ad Exposure (Instrumented Media)				0.0826
				(0.0657)
× Under-age				-0.0446
				(0.1000)
Mag Ad Exposure (Non-Ins Media)				0.0580***
				(0.0118)
× Under-age				-0.0044
				(0.0141)
TV Ad Exposure (Non-Ins Media)				0.0050***
				(0.0011)
× Under-age				-0.0016
				(0.0013)
Observations	6568	6568	6568	6568
R^2	0.084	0.099	0.619	0.088
F-Statistic (Mag Ad Exposure)				7.48
F-Statistic (× Under-age)				4.09
F-Statistic (TV Ad Exposure)				50.67
F-Statistic (× Under-age)				25.13

Robust standard errors in parentheses; All advertising measures divided by 100
All models include demographic controls, magazine issues read in the past 6 months, hours of TV viewed per week, and market (DMA) and survey wave (time) fixed effects.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Joint statitistical significance of under-age effect: $^{+}$ $p < 0.10$, $^{++}$ $p < 0.05$, $^{+++}$ $p < 0.01$

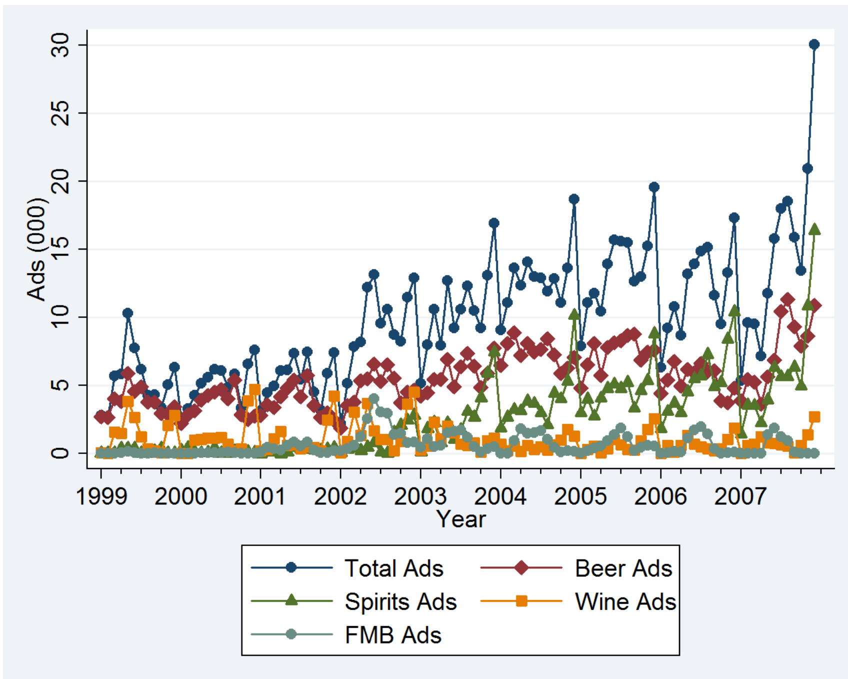

Figure 1.2: Number of alcohol advertisements, both in total and by alcohol type, on all television programs tracked by Kantar Media, 1999-2007. Note: FMB abbreviates flavored malt beverages. Source: Kantar Media.

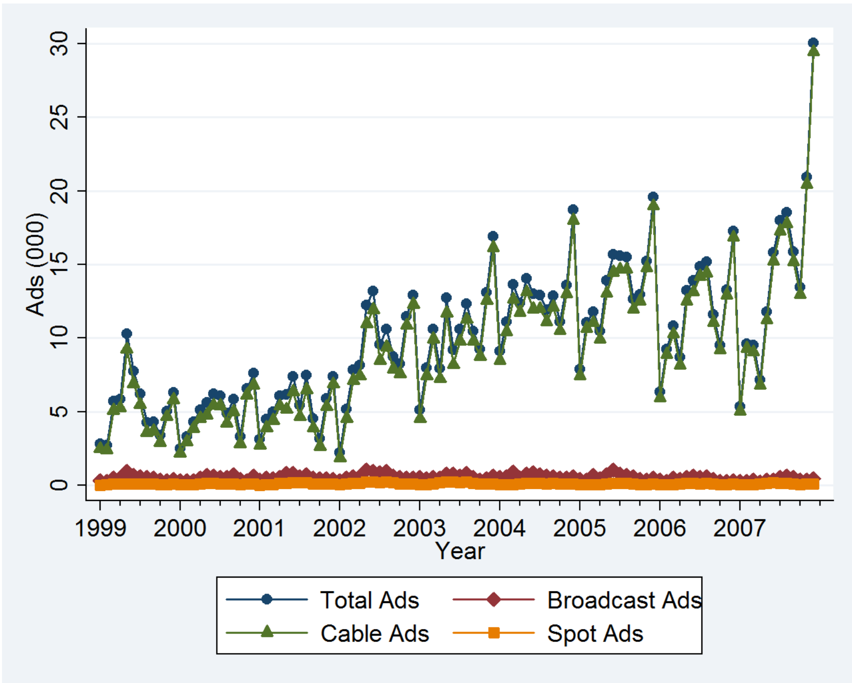

Figure 1.3: Number of alcohol advertisements, both in total and by television advertisement type, on all television programs tracked by Kantar Media, 1999-2007. Note: Spot ads are locally airing ads on broadcast networks and are scaled by the fraction of the U.S. population living in the local media market in which the ad appeared. Source: Kantar Media.

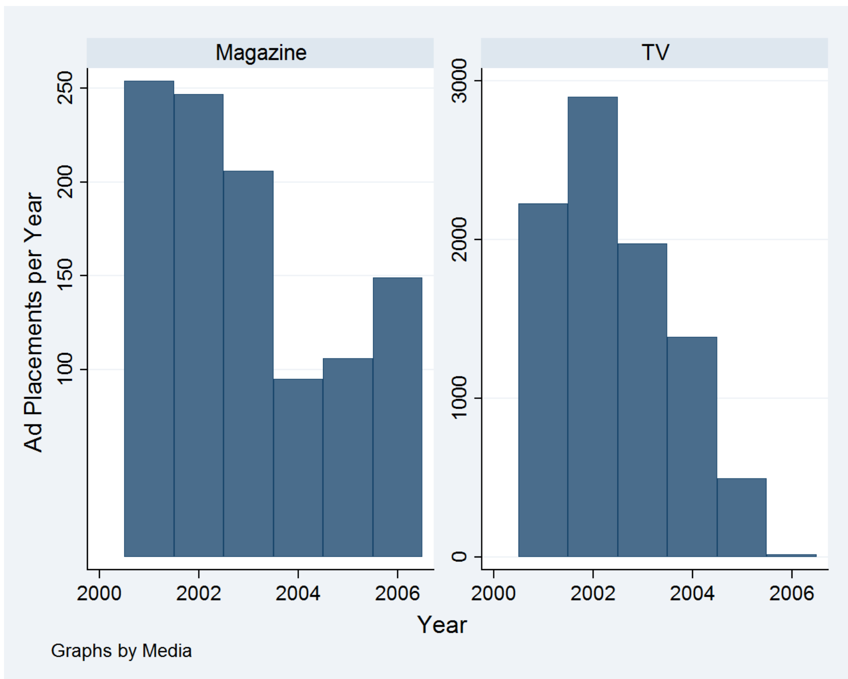

Figure 1.4: Number of alcohol advertisements appearing on the 8 television program and 5 magazines affected by the 2003 change in self-regulation governing the minimum allowed legal purchasing age audience. Source: Kantar Media.

Appendix A: Simmons National Consumer Survey

Survey Methods

The Simmons National Consumer Survey (NCS) is a nationally representative study measuring detailed product usage, media viewing, and demographics. Firms selling consumer goods, advertising agencies, and media companies are the main users of the NCS data.

The NCS is a repeated cross-section that is collected twice annually in spring and fall waves. Simmons uses a single-stage stratified sample covering the contiguous 48 states. Simmons splits the U.S. into 16 geographic regions (18 after the Spring 2003 Wave): the 14 (16) largest designated marketing areas (DMAs) and the balance of the contiguous U.S. split by the four Census Regions (Northeast, North Central, South, and West). Each of these regions is then split into four income strata. Simmons defines these strata as the set of telephone area code prefixes where a particular percentage of households have income over \$50,000[24] (66 percent or more, between 50 percent and 66 percent, between 33 percent and 50 percent, and 33 percent or less). This creates 64 (72) separate sampling strata. Simmons then samples households in these strata using random digit dialing, excluding area codes dedicated to mobile devices. The survey over samples higher income strata. A household in the highest income strata is four times more likely to be sampled than a household in the lowest strata.[25]

Simmons makes up to 16 attempts to contact selected households and two attempts to recontact initial refusers. Respondents receive cash incentives to participate. Simmons provides each person in the household age 6 and older with a Personal Booklet. The content of this booklet differs for the Kids (age 6 to 11), Teens (age 12 to 17), and Adults (age 18 and older) Surveys. This study mainly uses data from the Adults Survey, though the Teens Survey contributes to the magazine ratings measures. Each survey contains extensive questions in these areas: media usage (print, television, radio, and internet), demographics, lifestyle (interests, hobbies, etc.), product usage (alcohol, food, clothes, etc.), and retail shopping. The NCS is unique and extremely useful as it measures both media viewing and drinking.

[24] Simmons bases these strata on Census and Current Population Survey data.

[25] Households in the second and third highest income strata are three and twice as likely to be sampled respectively.

DMA Identification

The comprehensive advertising exposure measures used here require identifying the local television stations viewed by an individual. Though the NCS includes numerous geographic identifiers, it does not identify a DMA for every respondent. The survey directly identifies 12 DMAs in the Fall 2002 to Spring 2003 Waves and 14 DMAs in the Fall 2003 Wave to the Spring 2006 Wave. Respondents living in these 14 DMAs represent 88 percent of my sample. I use the state of residence and the general ranking of the DMA[26] to identify unique combinations of a state and DMA rank range. For example, Denver is the only DMA in Colorado, Wyoming, or Nebraska ranked from 11 to 25. Therefore, any respondent living in a DMA ranked 11 to 25 in those three states must live in the Denver DMA. This approach allows me to identify an additional 909 respondents in the sample. In total, the NCS and the DMA assignment algorithm identify 78 percent of NCS respondents age 18 to 24.

Alcohol Usage Measures

I use the NCS alcohol usage data to construct an indicator of any past 30 day drinking and the number of drinks consumed in that time. The NCS asks respondents to report the general number of drinks per month they consumed of each of 25 alcohol types (e.g., light domestic beer, vodka, and imported wine). The survey asks respondents to select the range of their consumption for each type: 0, 1 to 2, 3 to 4, 5 to 6, 7 to 10, 11 to 13, 14 to 19, 20 to 29, and 30 or more. For each response below the maximum, I assign an individual the midpoint value of the range they indicate. For the maximum, I assign individuals the maximum value plus half of the previous interval (35 drinks). I sum across all alcohol types to generate the total consumption in the last 30 days. Finally, I create an indicator of past 30 day alcohol use that equals one if one's total 30 day consumption is greater than zero.

The accuracy and validity of the alcohol usage measures in the NCS are particularly important for this study. While each respondent answers a personal survey, the NCS does not take particular steps to ensure responses are kept private from other household members. Youth concerned about parents or guardians seeing their responses may skip or alter their responses to the alcohol questions. I compare the NCS to the National Survey on Drug Use and Health (NSDUH) to study the validity of the NCS alcohol measures. The NSDUH is an annual survey of approximately 70,000 individuals age 12 and over. The survey is a major

[26]Nielsen ranks DMAs by the number of households with televisions. The NCS provides the range of the DMA rank for respondents that live in the top 100 DMAs (1 to 5, 5 to 10, 11 to 25, 26 to 50, and 51 to 100)

source of data on alcohol and illicit drug use, making it an ideal benchmark.

Table 1.6 compares 30 day drinking prevalence rates by age for the NCS and NSDUH. Since a primary concern is the effect of parental presence on reporting drinking, I also compare the prevalence rates for youth living in households with a parent present.[27] The results show that under-age NCS respondents are less likely to report drinking. Eighteen-year-old NCS respondents are particularly less likely to drinking. Surprisingly, NCS respondents age 19 and older living with or without a parent do not report drinking at different rates.

Table 1.6 compares past 30 day drinking volume for drinkers by age in the NCS and NSDUH. NCS respondents report drinking 7 fewer drinks on average, though this difference decreases with age. This difference may result from the different questions measuring drinking volume in the NSDUH. The survey asks respondents the number of the past 30 days they drank and the average amount they drank on one of those drinking days. Multiplying those two measures estimates the number of drinks consumed in the last 30 days, but the NCS and NSDUH methods may yield slightly different results.

Table 1.6 shows comparable "heavy" drinking prevalence rates in the NCS and NSDUH. In both surveys, the "heavy" drinking measure indicates whether a respondent drank 33 or more drinks in the past 30 days. Since these measures are based on the volume measures used in Table 1.6, the "heavy" drinking rates in both surveys show similar differences. Overall, NCS respondents are less likely to drink "heavily," which again may be due to the differences in how drinking volume is measured in each survey.

Figures 1.6, 1.6, and 1.6 compare trends in 30 day drinking prevalence, volume, and heavy drinking in both surveys for the sample period used in this study. These figures illustrate the general differences mentioned above: under-age prevalence and "heavy" drinking rates in the NCS are lower than the NSDUH and the average drinking volume in the NCS is lower than the NSDUH. Though the NCS measures appear to vary more, perhaps due to a smaller sample, both surveys show similar flat trends in prevalence, "heavy" drinking, and volume of alcohol consumption.

While the NCS does perfectly match the NSDUH or include measures of alcohol abuse, such as binge drinking, the NCS has the distinct advantage of rich data measuring media viewing. These data make the NCS particularly well suited for studying the effect of advertising on drinking.

[27] The NCS does not directly ask about parental presence. I define parental presence as living in a household with a "household head" older that 33 years.

Table 1.6: Comparison of NCS and NSDUH: 30 Day Drinking Prevalence

	NCS			NSDUH		
Age	All	Parent	No Parent	All	Parent	No Parent
18	0.30	0.29	0.39	0.42	0.41	0.56
19	0.38	0.37	0.44	0.50	0.47	0.62
20	0.46	0.45	0.50	0.55	0.51	0.69
21	0.62	0.64	0.58	0.68	0.64	0.79
22-24	0.65	0.66	0.62			
22-23				0.66	0.63	0.77

Source: 2000-2007 NCS and NSDUH.

Table 1.7: Comparison of NCS and NSDUH: Drinks Last 30 Days

	NCS			NSDUH		
Age	All	Parent	No Parent	All	Parent	No Parent
18	21.1	21.5	18.7	27.7	26.3	34.5
19	22.4	22.1	24.0	29.6	27.2	35.6
20	23.9	24.9	19.1	29.5	28.1	33.2
21	25.6	25.5	25.8	29.3	26.9	34.6
22-24	22.5	23.5	20.6			
22-23				27.7	26.5	32.2

Source: 2000-2007 NCS and NSDUH.

Table 1.8: Comparison of NCS and NSDUH: "Heavy" Drinking Last 30 Days

	NCS			NSDUH		
Age	All	Parent	No Parent	All	Parent	No Parent
18	0.21	0.22	0.15	0.35	0.33	0.40
19	0.21	0.21	0.22	0.36	0.34	0.42
20	0.25	0.27	0.17	0.36	0.35	0.40
21	0.26	0.26	0.24	0.35	0.32	0.41
22-24	0.22	0.23	0.20			
22-23				0.33	0.32	0.38

Source: 2000-2007 NCS and NSDUH.

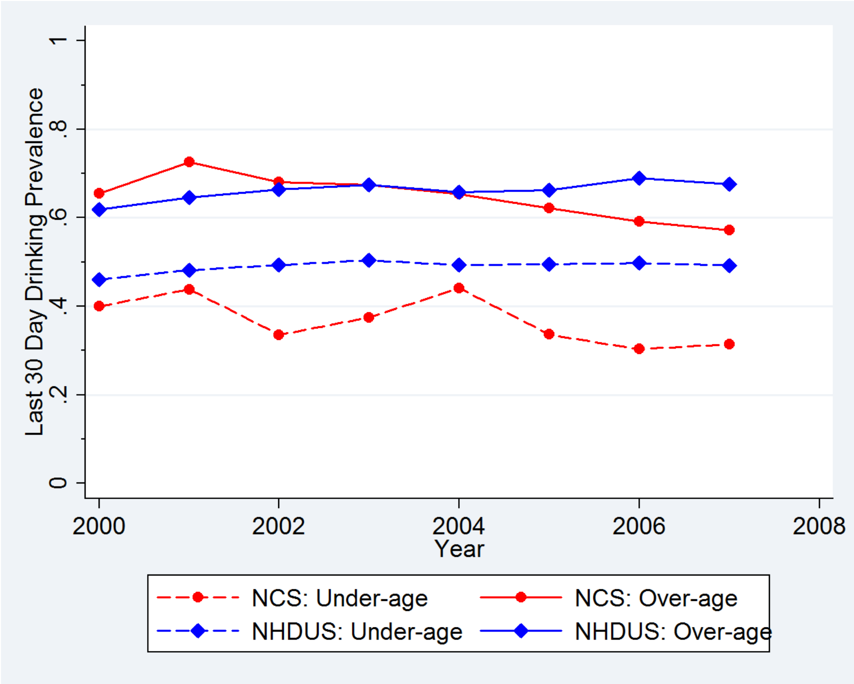

Figure 1.5: Comparison of NCS and NSDUH: 30 Day Drinking Prevalence. Source: 2000-2007 NCS and NSDUH.

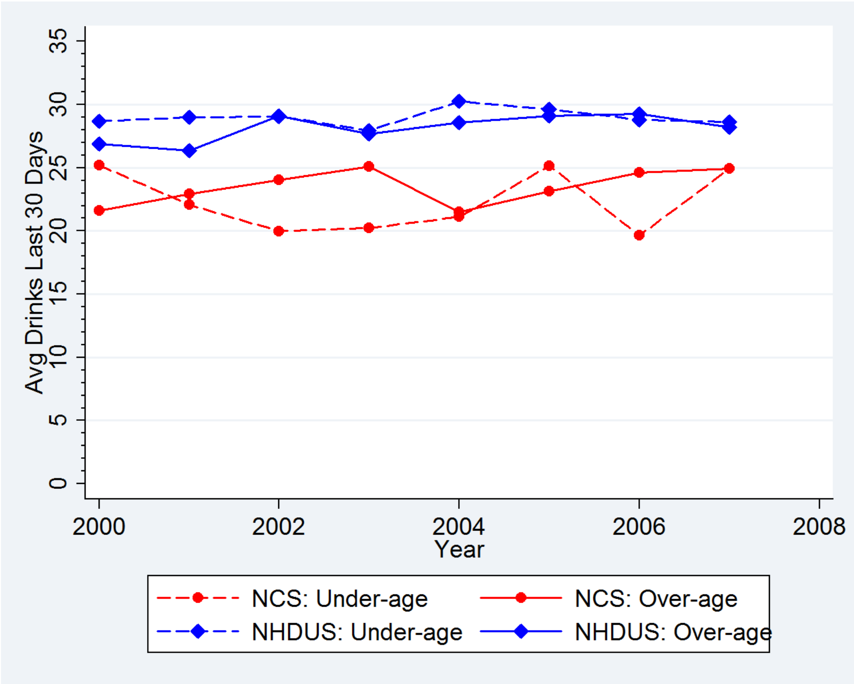

Figure 1.6: Comparison of NCS and NSDUH: Drinks Last 30 Days. Source: 2000-2007 NCS and NSDUH.

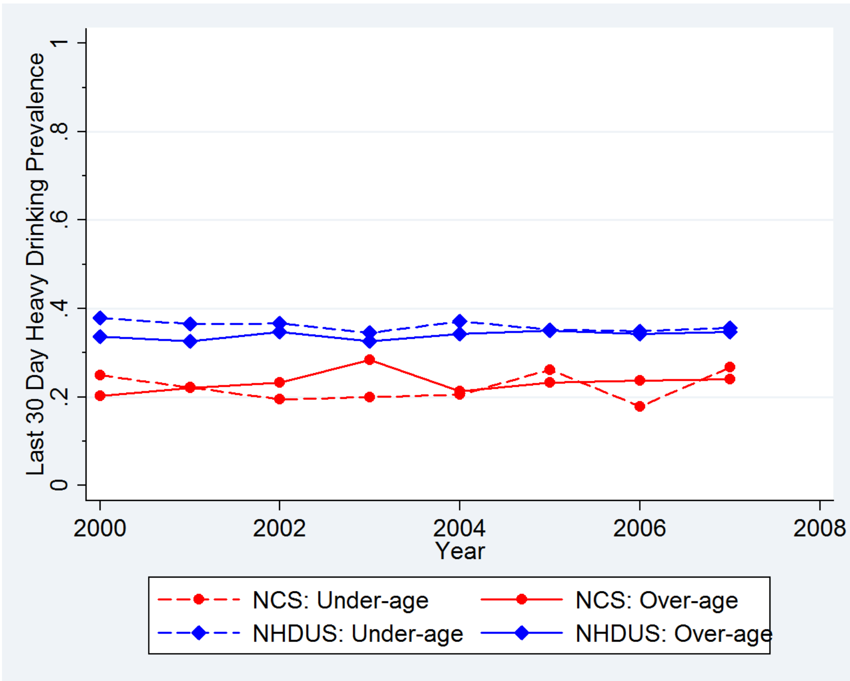

Figure 1.7: Comparison of NCS and NSDUH: "Heavy" Drinking Last 30 Days. Source: 2000-2007 NCS and NSDUH.

Appendix B: Media Exposure Estimates

Magazine Advertising Exposure

The magazine advertising exposure measure follows Avery et al (2007a) and uses data on each individual's magazine reading practices. Each NCS respondent indicates the fraction of the last four issues he read of each of 182 magazines. I assume that this fraction proxies for the fraction of issues of each magazine the respondent has read over the past six months (denoted as $view_{imt}$), where the subscripts denote respondent i, magazine m, wave t. I multiply the fraction of issues of each magazine the respondent read times the number of alcohol advertisements that appeared in magazine m. I then sum across all magazines an individual reads. Potential exposure to alcohol advertisements of respondent i in wave t is given by:

$$ad_{it}^{mag} = \sum_{m} view_{imt} \times ads_{mt} \tag{1.6}$$

Television Advertising Exposure

I adopt a similar method to construct each individual's exposure to alcohol advertising on television; however, there are some differences. The key difference is that the NCS includes many more television programs and the survey questions about these programs differ based on program type. The NCS provides data on the viewing of nationally broadcast network television programs, cable television programs, regular season sports, and special programs that air infrequently or only once. The Kantar advertising data measure the number of alcohol advertisements that aired during each program over the whole sample period on all of these media. For broadcast network programming, the Kantar data also indicate whether the ad aired nationally or locally, and, if locally, the DMAs in which the ad appeared. As with print media, I use individual reports on what television each person watches and when. NCS respondents report how many of the last few episodes (four episodes for weekly shows, five episodes for daily shows) of regularly occurring major broadcast network programs (NBC, ABC, CBS, FOX, WB/CW, UPN/MY) that they watched. The survey also asks when respondents last watched (if ever) specific regularly occurring cable programs. Respondents also indicate how often they watched regular season sports in the past year. Finally, respondents report whether they watched special programming the last time it aired. I use this information as follows:

Nationally broadcast television programs

I match alcohol advertising that appears on nationally broadcast television programs to programs included in the NCS. I generate a probability that a respondent viewed a specific airing of each program. This probability is equal to the number of previous airings the respondent reported viewing divided by the number they could have viewed. For example, if a respondent reports watching three of the last four episodes of *Seinfeld*, I assign her a probability of .75 that she viewed any one of the last four episodes. I then assume that this probability applies to every airing of that program in the previous six months. For each program a person views, I multiply the number of ads on that program in the six months prior to the survey times the fraction of airings she views. I sum this measure over the all broadcast programs to generate an estimate of exposure to alcohol ads on all broadcast programs.

Regular Sports

The NCS includes questions on how often one watches the regular season sports. The questions refer to particular sports on a specific networks (e.g., *CBS College Football and ESPN MLB Baseball* and to more generic sports titles (e.g., *Bowling* and *MLS Soccer*). Similar to broadcast programs, I create the fraction of airings a person reports viewing based on how many times he says he watched the sport in the last year. For each sport, I multiply the number of ads matched to that sports program from the Kantar data times the fraction of airings a person reports viewing to create their total exposure on that sports program. I then sum this measure over all regular sports programming to create the total exposure to alcohol advertising on regular season sports.

Cable television programs

I use a slightly different method to match advertisements aired on cable television programs because the NCS collects different information on cable television viewing and cable networks rerun their programming more often. To address repeats, I use the Kantar data to determine how many times a program aired in a given day. Since I cannot know which airing an individual viewed, I divide the number of ads appearing on every airing of that program by the number of times the program aired that day to create the average number of ads per airing for that day. This approach assumes people are more likely to view one airing of a program a day than watch several repeats. The NCS asks respondents whether they watch each of approximately five programs for each cable network. The survey tells respondents to indicate whether they watched each program in the last four weeks and if they watched the program in the last seven days. If the respondent reports viewing a program in the last seven days I assume he watch the program every day it aired. If the respondent reports watching the program in the last four weeks but not in the last seven days, then I assume he watches the program half the days it airs. Similar to broadcast programs, I multiply the six month total of the daily average ads on a program times the fraction of days a person watches to estimate the number of ads he saw on that program. I then sum these estimates over all programs a person reports viewing.

Single-event specials

I use a similar procedure to count advertisements placed on single-airing specials (e.g., the Academy Awards) and special sporting events (e.g., the World Series). The NCS asks respondents if they viewed each of these special programs the last time they aired. If a person says he watched a special program, I assume he watched the event in its entirety (e.g., every game of the World Series). I include in his exposure measure all alcohol advertisements that aired in his DMA during the special programs he saw.

Taken together, I use all of the above information to measure each respondent's exposure to television alcohol advertising. The television exposure measure is formally constructed as the sum of the advertisements that appeared on each program a person said he watches. Thus potential exposure to alcohol television advertisements of respondent i in year t and market m is given by:

$$ad_{it}^{tv} = \sum_{n} view_{imnt} \times ads_{mnt} + \sum_{c} view_{ict} \times ads_{ct} + \sum_{r} view_{imrt} \times ads_{mrt} + \sum_{s} watch_{imst} \times ads_{mst} \tag{1.7}$$

where subscripts refer to the following: n refers to each nationally broadcast television show; c refers to each cable television show; r refers to each regular season sports show; and s refers to special shows that aired only one time in a given year. $view_{imxt}$ measures the fraction of airings of program x viewed by person i. $watch_{imst}$ measures whether person i watched special s the last time it aired.

Comparison Advertising of Television Exposure Measures

In Figure 1.6, I compare the average exposure of NCS respondents to television alcohol advertising with estimated average exposure to ads as computed by Nielsen from Gross Ratings Points (GRP)(CAMY, 2010a). GRPs measure the average number of advertisements a person in a specific demographic group saw in a year. The dashed lines plot Nielsen's Gross Ratings Points for youth ages 12 to 20 (triangles) and adults ages 21 to 34 (diamonds). The solid line plots estimated exposure for my sample (age 18 to 24) in each year using my NCS data and exposure measure. The exposure measure used in my study tracks yearly trends in exposure closely. While the NCS estimated exposure is, on average, 4.4 times as large as Nielsen's measure of 21 to 34 year old exposure and 6.6 times as large as the measure of 12 to 20 year old exposure, these difference likely arise because I use NCS time diary questions about television viewing while Nielsen more directly measures what a person watches. Nielsen uses in-home television tracking boxes and can therefore determine exactly which airing of a program an individual views. Though the Nielsen GRP is an average of exposure of all people in the age group in a given year or in a given year and media market, the NCS exposure estimates measure each individual's exposure. Importantly, unlike the Nielsen GRP measure, the NCS measure does not assign the same level of exposure to every person in a given DMA. Instead I use variation in exposure that arises because individuals differ in their tastes, family composition, and work schedules. These differences cause people to watch different programs at different times. This individual variation in advertising exposure is vital when studying individual consumption behavior. Furthermore, since my measures likely overestimate exposure for all individuals, this difference will be fully captured by the size of the coefficient estimate of the effect of advertising. If the Nielsen GRP measures are taken as truth, scaling my coefficients by a factor of 4.4 to 6.6 would make them comparable to estimates using GRP advertising exposure measures. To further simplify the interpretation of my results, I relate my coefficient estimates in terms of percentage changes in average advertising exposure, which makes scaling unnecessary.

Figure 1.6 compares the average magazine advertising exposure of NCS respondents to similar GRP measures created using Mediamark Research Inc. (MRI) magazine ratings data and Kantar Media[28] advertising data (Center on Alcohol Marketing and Youth, 2010*b*). Because MRI collects magazine ratings data using a survey similar to the NCS, the NCS magazine measure I use in this study is directly comparable to the independent MRI measure.

[28] Kantar Media was previously known as TNS Media.

Figure 1.6 shows that the NCS estimate measures a sharper decline in advertising exposure over the five year period. The average exposure of the NCS sample I use in this study is most comparable to the 12 to 20 year old MRI GRP measure. Because my magazine exposure measure is close in magnitude to independent measures used in other studies, the magazine advertising effect coefficient estimates found in this paper should be directly comparable.

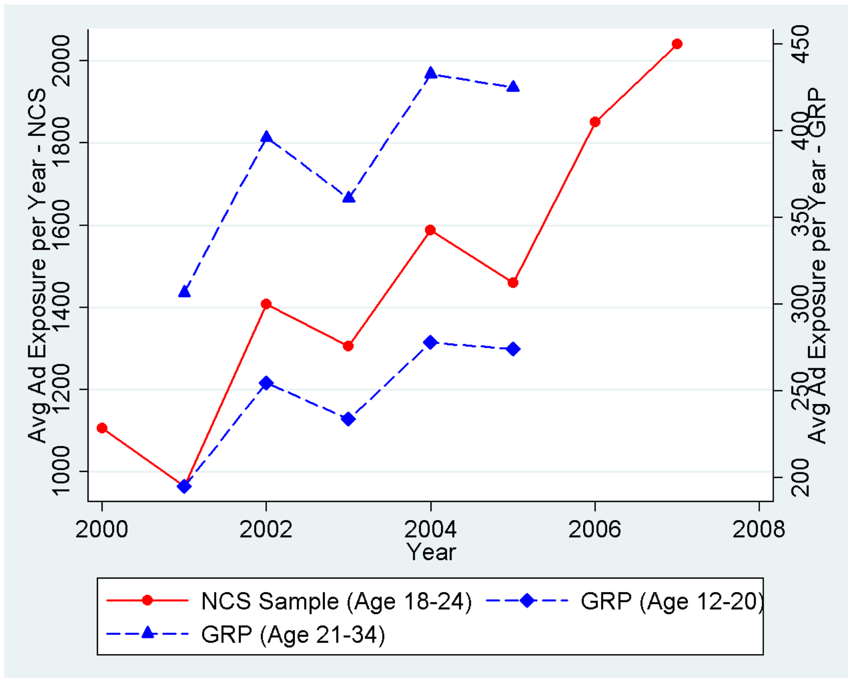

Figure 1.8: Comparison of NCS and Nielsen GRP Exposure Measures: Television Advertising. Source: 2002-2006 NCS, Kantar Media, and Center on Alcohol Marketing and Youth.

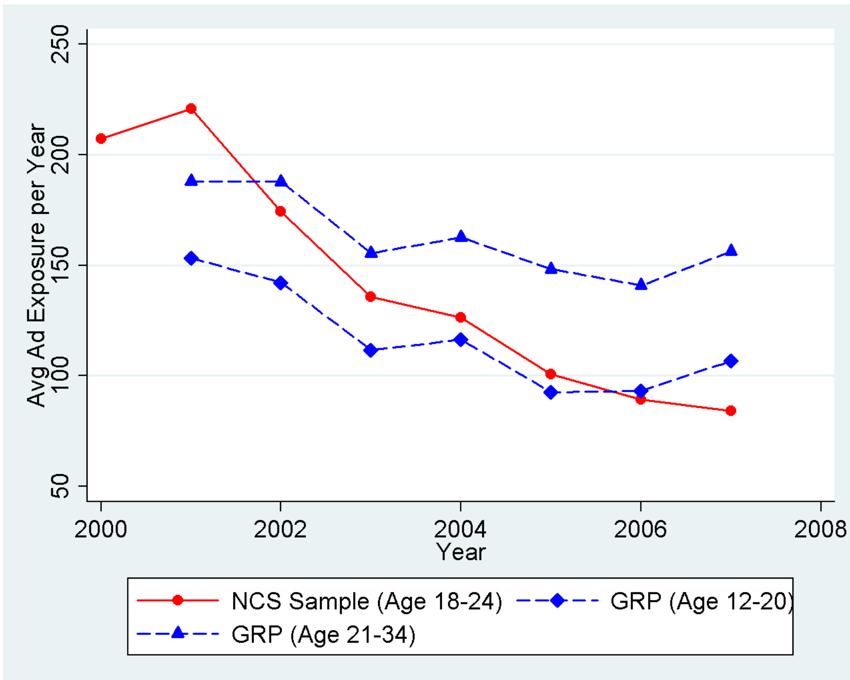

Figure 1.9: Comparison of NCS and Nielsen GRP Exposure Measures: Magazine Advertising. Source: 2002-2006 NCS, Kantar Media, and Center on Alcohol Marketing and Youth.

Appendix C: IV Analysis

Comparison of Treated and Untreated & First-Stage Results

Table 1.9: Comparision of Instrument Media Viewers and Non-Viewers

	Non-Viewers		Viewers	
	Mean	SD	Mean	SD
Any Drinks in L30 Days	0.47		0.56	
Drinks in L30 Days	8.92	17.38	14.89	24.13
Heavy Drinking	0.08		0.15	
Mag Ad Exposure	0.32	0.55	1.08	1.25
Mag Ad Exposure (Instrumented Media)	0.00	0.00	0.09	0.16
TV Ad Exposure	3.77	5.98	10.76	13.57
TV Ad Exposure (Instrumented Media)	0.00	0.00	0.37	0.61
Female	0.53		0.56	
Educ: Less than HS	0.21		0.16	
Educ: Some College	0.30		0.37	
Educ: College	0.13		0.12	
Full-Time Student	0.23		0.29	
Part-Time Student	0.06		0.08	
Employed Full-Time	0.43		0.41	
Employed Part-Time	0.24		0.27	
Married	0.47		0.32	
Age	21.04	1.98	20.66	1.99
Under-age	0.41		0.50	
Hispanic	0.48		0.36	
Black	0.03		0.12	
Other Race	0.18		0.16	
Number of Adults in HH	3.42	1.29	3.30	1.17
Real HH Income (000)	82.96	65.63	90.69	66.48
Likely Parent in HH	0.70		0.79	
Spanish Predom Lang in HH	0.33		0.17	
Other Lang Predom in HH	0.04		0.03	
Mag Issues Read	56.46	64.33	143.03	145.72
Total Count of Netwk Shows	10.03	13.54	21.92	33.50
Hours spent viewing netwk TV per week, sum of daypts	14.12	15.19	17.38	17.02
Total Hours viewing cable TV per week	8.33	13.94	21.83	29.30
Total Count of cable Shows watched	7.57	9.96	25.54	39.68
Read No Magazines	0.23		0.03	
Watch No TV	0.07		0.01	
Observations	6390		6362	

Table 1.10: Comparision of Instrument Media Viewers Pre and Post Rule Change

	2002-2003		2004-2006	
	Mean	SD	Mean	SD
Any Drinks in L30 Days	0.58		0.55	
Drinks in L30 Days	14.65	23.21	15.08	24.87
Heavy Drinking	0.15		0.16	
Mag Ad Exposure	1.26	1.29	0.93	1.19
Mag Ad Exposure (Instrumented Media)	0.10	0.18	0.09	0.15
TV Ad Exposure	7.91	7.61	13.14	16.67
TV Ad Exposure (Instrumented Media)	0.62	0.74	0.16	0.36
Female	0.57		0.56	
Educ: Less than HS	0.12		0.20	
Educ: Some College	0.46		0.30	
Educ: College	0.14		0.10	
Full-Time Student	0.32		0.27	
Part-Time Student	0.08		0.08	
Employed Full-Time	0.44		0.39	
Employed Part-Time	0.28		0.26	
Married	0.10		0.50	
Age	20.73	1.97	20.61	2.01
Under-age	0.48		0.51	
Hispanic	0.15		0.53	
Black	0.13		0.11	
Other Race	0.07		0.24	
Number of Adults in HH	3.14	1.08	3.43	1.22
Real HH Income (000)	101.40	68.51	81.71	63.36
Likely Parent in HH	0.80		0.78	
Spanish Predom Lang in HH	0.02		0.29	
Other Lang Predom in HH	0.04		0.03	
Mag Issues Read	142.11	120.74	143.80	163.78
Total Count of Netwk Shows	20.49	24.90	23.12	39.25
Hours spent viewing netwk TV per week, sum of daypts	16.26	16.42	18.33	17.46
Total Hours viewing cable TV per week	21.94	25.39	21.74	32.22
Total Count of cable Shows watched	22.07	26.53	28.45	47.81
Read No Magazines	0.02		0.04	
Watch No TV	0.01		0.01	
Observations	2902		3460	

Table 1.11: First Stage Estimates of Alcohol Advertising Exposure - Prevalence

	(1) Ins Mag Ad Exp	(2) × Uage	(3) Ins TV Ad Exp	(4) × Uage
106 & Park	0.0231***	0.0013*	0.0115	0.0142***
	(0.0057)	(0.0008)	(0.0212)	(0.0044)
Behind the Music	-0.0029	0.0015*	-0.4248***	-0.0047
	(0.0047)	(0.0008)	(0.0304)	(0.0033)
BET Comic View	-0.0073	0.0002	-0.3475***	0.0018
	(0.0063)	(0.0008)	(0.0268)	(0.0044)
Crank Yankers	0.0109*	-0.0040***	-0.0989**	-0.0063
	(0.0064)	(0.0013)	(0.0412)	(0.0054)
Croc Hunter	-0.0061*	0.0032***	-0.1225***	0.0053**
	(0.0037)	(0.0006)	(0.0165)	(0.0026)
Driven	0.0124**	-0.0056***	-0.5383***	0.0055
	(0.0059)	(0.0011)	(0.0490)	(0.0046)
Rap City	-0.0112*	0.0009	0.0494**	-0.0197***
	(0.0063)	(0.0008)	(0.0229)	(0.0045)
BET Top 25 Countdown	-0.0109	0.0011	-0.1744***	0.0154***
	(0.0086)	(0.0011)	(0.0309)	(0.0049)
106 & Park × Underage	-0.0099	0.0114	-0.0060	-0.0108
	(0.0089)	(0.0070)	(0.0301)	(0.0226)
Behind the Music × Underage	0.0022	-0.0042	-0.0024	-0.4131***
	(0.0066)	(0.0047)	(0.0416)	(0.0302)
BET Comic View × Underage	0.0111	0.0035	0.0529	-0.2989***
	(0.0090)	(0.0067)	(0.0366)	(0.0265)
Crank Yankers × Underage	-0.0114	0.0069	-0.0817	-0.1725***
	(0.0108)	(0.0090)	(0.0590)	(0.0448)
Croc Hunter × Underage	-0.0146**	-0.0265***	-0.0087	-0.1432***
	(0.0071)	(0.0062)	(0.0260)	(0.0217)
Driven × Underage	-0.0153*	0.0065	0.0301	-0.5174***
	(0.0087)	(0.0065)	(0.0704)	(0.0515)
Rap City × Underage	0.0118	-0.0010	0.0576*	0.1399***
	(0.0085)	(0.0060)	(0.0309)	(0.0223)
BET Top 25 Countdown × Underage	0.0046	-0.0089	-0.0314	-0.2444***
	(0.0117)	(0.0082)	(0.0422)	(0.0304)
Allure	-0.0013	0.0049***	-0.0025	-0.0050
	(0.0072)	(0.0012)	(0.0195)	(0.0046)
Automobile	0.0149	-0.0021	-0.0330	-0.0119**
	(0.0126)	(0.0016)	(0.0287)	(0.0051)
ESPN	-0.0573***	-0.0042***	0.0354	-0.0022
	(0.0213)	(0.0012)	(0.0243)	(0.0039)
Spin	-0.1547***	-0.0009	-0.0252	-0.0014
	(0.0152)	(0.0011)	(0.0278)	(0.0051)
Vibe	-0.0101	0.0038***	0.0144	-0.0034
	(0.0136)	(0.0011)	(0.0287)	(0.0053)
Allure × Underage	-0.0033	-0.0139**	-0.0098	-0.0043
	(0.0100)	(0.0070)	(0.0295)	(0.0236)
Automobile × Underage	0.0081	0.0275***	0.0495	0.0359
	(0.0162)	(0.0101)	(0.0390)	(0.0270)
ESPN × Underage	-0.0041	-0.0552***	-0.0337	0.0005
	(0.0292)	(0.0206)	(0.0324)	(0.0226)
Spin × Underage	0.0404	-0.1127***	0.0342	0.0205
	(0.0261)	(0.0219)	(0.0509)	(0.0433)
Vibe × Underage	-0.0002	-0.0172	-0.0742*	-0.0598**
	(0.0190)	(0.0135)	(0.0395)	(0.0289)
Observations	12752	12752	12752	12752
R^2	0.919	0.921	0.926	0.924
F-Statistic on Instruments	8.88	5.72	82.04	48.40
Partial R^2 on Instruments	0.07	0.05	0.36	0.35

Robust standard errors in parentheses; All advertising measures divided by 100
All models include demographic controls, magazine issues read in the past 6 months,
hours of TV viewed per week, and market (DMA) and survey wave (time) fixed effects.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.12: First Stage Estimates of Alcohol Advertising Exposure - Volume

	(1) Ins Mag Ad Exp	(2) × Uage	(3) Ins TV Ad Exp	(4) × Uage
106 & Park	0.0251***	0.0002	0.0118	0.0057
	(0.0069)	(0.0007)	(0.0273)	(0.0046)
Behind the Music	-0.0003	0.0005	-0.4132***	-0.0043
	(0.0052)	(0.0009)	(0.0346)	(0.0037)
BET Comic View	-0.0100	-0.0003	-0.3418***	0.0053
	(0.0074)	(0.0007)	(0.0327)	(0.0044)
Crank Yankers	0.0059	-0.0034***	-0.0948**	-0.0100*
	(0.0070)	(0.0013)	(0.0441)	(0.0054)
Croc Hunter	-0.0068	0.0017**	-0.1182***	0.0039
	(0.0047)	(0.0007)	(0.0202)	(0.0029)
Driven	0.0143**	-0.0050***	-0.5600***	-0.0009
	(0.0067)	(0.0013)	(0.0558)	(0.0047)
Rap City	-0.0117	0.0007	0.0447	-0.0130***
	(0.0077)	(0.0008)	(0.0295)	(0.0047)
BET Top 25 Countdown	-0.0146	0.0002	-0.1737***	0.0062
	(0.0100)	(0.0011)	(0.0393)	(0.0048)
106 & Park × Underage	-0.0152	0.0083	-0.0056	-0.0033
	(0.0115)	(0.0094)	(0.0430)	(0.0356)
Behind the Music × Underage	0.0081	0.0037	-0.0055	-0.3997***
	(0.0078)	(0.0062)	(0.0547)	(0.0465)
BET Comic View × Underage	0.0064	-0.0042	0.0159	-0.3345***
	(0.0124)	(0.0102)	(0.0500)	(0.0404)
Crank Yankers × Underage	0.0128	0.0297**	-0.0416	-0.1122*
	(0.0132)	(0.0118)	(0.0754)	(0.0672)
Croc Hunter × Underage	-0.0156	-0.0271***	-0.0363	-0.1705***
	(0.0100)	(0.0090)	(0.0374)	(0.0337)
Driven × Underage	-0.0174	0.0062	-0.0263	-0.5902***
	(0.0111)	(0.0090)	(0.0888)	(0.0714)
Rap City × Underage	0.0209**	0.0077	0.0375	0.1149***
	(0.0106)	(0.0076)	(0.0418)	(0.0318)
BET Top 25 Countdown × Underage	0.0062	-0.0110	-0.0156	-0.2269***
	(0.0152)	(0.0115)	(0.0611)	(0.0496)
Allure	-0.0068	0.0026**	-0.0066	0.0016
	(0.0092)	(0.0012)	(0.0261)	(0.0049)
Automobile	0.0252	-0.0005	-0.0210	-0.0061
	(0.0155)	(0.0015)	(0.0324)	(0.0062)
ESPN	-0.0839***	-0.0025**	0.0023	-0.0056
	(0.0248)	(0.0012)	(0.0295)	(0.0051)
Spin	-0.1526***	-0.0011	-0.0138	0.0014
	(0.0180)	(0.0011)	(0.0323)	(0.0049)
Vibe	-0.0006	0.0016	0.0310	0.0021
	(0.0159)	(0.0011)	(0.0368)	(0.0054)
Allure × Underage	-0.0097	-0.0253*	0.0028	0.0068
	(0.0170)	(0.0144)	(0.0422)	(0.0356)
Automobile × Underage	-0.0211	0.0100	0.0219	0.0109
	(0.0190)	(0.0112)	(0.0482)	(0.0361)
ESPN × Underage	-0.0831**	-0.1662***	-0.0595	-0.0519
	(0.0393)	(0.0314)	(0.0441)	(0.0346)
Spin × Underage	0.0255	-0.1260***	0.0239	0.0123
	(0.0316)	(0.0272)	(0.0625)	(0.0553)
Vibe × Underage	0.0182	0.0142	-0.0375	-0.0284
	(0.0230)	(0.0171)	(0.0545)	(0.0422)
Observations	6568	6568	6568	6568
R^2	0.929	0.943	0.930	0.934
F-Statistic on Instruments	7.48	4.09	50.67	25.13
Partial R^2 on Instruments	0.14	0.17	0.38	0.40

Robust standard errors in parentheses; All advertising measures divided by 100
All models include demographic controls, magazine issues read in the past 6 months, hours of TV viewed per week, and market (DMA) and survey wave (time) fixed effects.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

CHAPTER 2

ADOPTION OF TELEVISION ADVERTISING FOLLOWING THE UNRAVELING OF THE BAN ON TELEVISION SPIRITS ADVERTISING

This research is supported by the Bronfenbrenner Life Course Center Innovative Research Program and award 1R01-AA019777-01A1 from the National Institute on Alcohol Abuse and Alcoholism.

2.1 Introduction

While alcohol is enjoyed responsibly and legally by the majority of drinkers, alcohol abuse, particularly among young drinkers, is a serious public health concern. 36.2 percent of 18- to 20-year olds and 46.1 percent of 21- to 25-year olds report heavy episodic drinking, defined as consuming 5 or more drinks on one occasion (SAMSHA, 2007). Heavy drinking causes a number of fatal and non-fatal consequences. The CDC estimates that there were 75,766 alcohol-attributable deaths in 2004 (CDC, 2004), making alcohol abuse the third leading cause of preventable death in the U.S. (Mokdad et al., 2004). Alcohol is also implicated in college sexual assaults and injuries (Hingson et al., 2005), crime (Carpenter, 2007), and the spread of sexually transmitted diseases (Chesson, Harrison and Kassler, 2000). These consequences include significant negative externalities, costs alcohol abusers impose on others (Kenkel, 1993). The estimated cost of excessive drinking in the U.S. was $223.5 billion in 2006 (Bouchery et al., 2011).

The costs of alcohol abuse focus attention on the significant sums spent marketing alcohol and the potential influence of advertising on drinking. In 2005, the alcohol industry spent $2 billion advertising its products. The alcohol industry argues that advertising influences the consumers' brand choices but does not affect whether or how much they drink. If this is true, alcohol advertising does not threaten the public health. However, there is considerable disagreement among researchers over whether advertising drives people to drink and drink more. A review by Hastings et al. (2005) finds convincing evidence in several studies that advertising causes youth to drink. Nelson (2010) reviews largely the same literature and finds that shortcomings of previous studies do not allow for the interpretation of a causal effect of advertising on drinking. Molloy (2012) addresses some of these shortcomings and finds little evidence of an effect of advertising on youth drinking prevalence or volume. However, the grave consequences of alcohol abuse warrants further study of alcohol advertising.

Despite the paucity of evidence of a causal effect of advertising on drinking, alcohol firms have established significant advertising self-regulation. The most substantial self-imposed rule was a 1933 agreement by spirits manufacturers to ban radio advertising of their products. In 1948, this agreement was extended to include television advertising. This self-imposed ban lasted 63 years until 1996, when the Distilled Spirits Council of the United States (DISCUS), the trade group representing the spirits industry, agreed unanimously to abandon the ban. The result was a short period of relative inactivity, with $35 million spent on spirits television alcohol advertising from 1996-2001, followed by large year-on-year increases in television alcohol advertising reaching $115 million in yearly spending in 2007 (Kantar Media 1995–2010).

The lifting of the broadcast ban on distilled spirits presents an interesting case study in firm advertising behavior. The sudden, novel access of a mature industry to a well-developed advertising medium is a rare occurrence. Understanding how firms react to the lifting of the ban can provide insight into the role advertising plays in markets. The timing and level of the use of television advertising differ significantly across different types of spirits. Cordial brands began advertising on television in 2001, while Tequila brands did not begin sizable spending on television until 2004. There are also large differences in the adoption of television advertising by different brand within a type of spirits. *Smirnoff* vodka begins advertising on television in 2003 and spends 92% of its advertising spending on TV by 2007. *Grey Goose* vodka also begins TV advertising in 2003, but by 2007 TV advertising only comprised 10% of its advertising spending. Differing substantially from these brands, *Skyy* vodka never advertises on television and continued spending millions of dollars each year on magazine advertising.

This paper studies spirits brand advertising choices following the end of the 1996 self-imposed broadcast advertising ban. I use novel data that are particularly well-suited to exploring firm advertising behavior. I model several measures of advertising adoption: the start of television advertising, defined as spending more that $500,000 on TV advertising in a year; the switch to television advertising, defined as spending 50% or more of a brand's advertising budget on TV; and whether the brand is the "first mover" in its market.[1] I investigate how these behaviors are related to the demographic characteristics, media consumption, and drinking behavior of consumers of the brand; market concentration; the brand's power in the market; and the advertising behavior of other brands in the market. I do not intend to estimate the causal effect of any of these factors on television advertising. However, sev-

[1] Market is defined by spirits type (e.g., vodka, rum).

eral of my models estimate the relationship between market and brand characteristics prior to the ban and post-ban firm behavior, eliminating the threat of simultaneity to a causal interpretation of my results. Generally, I perform a "sophisticated description" of the correlates with television advertising adoption, shedding some light on factors that may drive the switch to TV for some brands.

I find the strongest correlate with television adoption is competitive pressure on a market leader by competitors. Brands that lead their spirits type in market share are significantly more likely to advertise on television if their closest competitors have larger market share. However, I find little evidence that the characteristics of a brand's consumers, including their age, gender, race, income, education, and magazine reading and television viewing habits, are related to the decision to move to television advertising. Although, brands that are more popular with beer drinkers are less likely to start television advertising and brands more popular with wine drinkers are more likely to move to television sooner. Overall, these results suggest that spirits advertising may largely divide market shares between competitors rather than grow the size of a market.

The rest of the paper is organized as follows: Section 2 provides background on the spirits industry and the unraveling of its television advertising ban. I also review several theories concerning advertising's role in markets. Section 3 and 4 describe my data and econometric methods. Section 5 reports summary statistics and model results. Section 6 discusses the results and concludes.

2.2 Background

Advertising: Basic Theory

Economic theory has largely struggled to explain the existence of advertising. A rational consumer with perfect information would not be expected to be swayed by 30 second commercials advocating the merits of a particular brand of beer. Three main theories have emerged to explain why advertising might affect consumer behavior (Bagwell, 2007). The "persuasive view" posits that advertising directly affects consumers' preferences, creating spurious product differentiation and leading them to prefer one product over an identical good of another brand. The "informative view" states that advertising is a convenient source of product information that would otherwise be more costly to obtain. Nelson (1970) presents evidence that advertising for "experience goods," those whose quality must be experienced through consumption, provides valuable information about product quality. Nelson (1974) provides the further insight that the ratio of television advertising to magazine advertising is higher for "experience goods." He argues that this is because advertisements for "experience goods" provide "soft", or imperfect, information about product quality, which is less useful to consumers than "hard," objective information about product characteristics. Because consumers value "soft" information less, they must be forced to watch such advertisements on television as they are less willing to voluntarily read magazine advertisements. The "complementary" view theorizes that consuming a product and viewing its advertisements are complements. Stigler and Becker (1977) create a model with stable preferences that implies that seeing more advertising for a product raise the marginal utility of consuming that product. These views suggest very different mechanisms for how advertising affects consumer behavior.

While the three main theoretical views of how advertising affects consumer demand are quite different, empirically testing which view or views applies in a particular market is very difficult. The "informative view" implies that a firms should target advertising at consumers that have high expected utility from consuming their product but little information about its quality. If a vodka brand believes their advertising is informative and vodka is an "experience good", they should target vodka drinkers that do not consume their brand and have not seen a substantial amount of advertising for that brand. This implies a post-ban strategy of advertising on television if a large fraction of other-brand vodka drinkers watch a large amount of television but read few magazines. However, the "persuasive" and "complementary" views can also explain this strategy if there are strongly decreasing marginal returns to advertising. That is, if the first few ads a person views are the most "persuasive" or "complementary," then a post-ban strategy of targeting those who favor television over magazines may also have a high return. Though these three theories are informative, they provide scant testable hypotheses for which alcohol brands may move to television advertising.

Ban of Broadcast Advertising of Spirits

The distilled spirits industry agreed to a voluntary ban on radio advertising in 1936, only three years after the Twenty-First amendment repealed Prohibition (Hemphill, 1998). The industry extended the ban to include television in 1948. Spirits firms almost universally adhered to this agreement for 60 years. However, by the mid-1990's spirits had lost a considerable share of alcohol consumption to wine and beer, falling from 44 percent in 1970 to 29 percent in 1995 (Hemphill, 1998). Seagram unilaterally violated the ban in June 1996, airing a commercial in Corpus Christi, Texas for Crown Royal Canadian whisky. Several other spirits firms followed Seagram, either airing television advertisements or making plans to do so. The decision to violate the ban faced immediate criticism from President Clinton, Members of Congress, and private groups such as Mothers Against Drunk Driving and the Center for Science in the Public Interest. On November 7, the policy board of DISCUS voted unanimously to abandon the broadcast advertising ban. This decision was met with further disapproval from politicians and the public, a campaign to legislate the ban, an investigation by the Federal Trade Commission, and the threat of an investigation by the Federal Communication Commission. However, the May 1996 ruling of the Supreme Court in *44 Liquormart, Inc. v. Rhode Island*[2] affirming the right of liquor stores to advertise prices established a strong precedent for the First Amendment right to commercial speech. This ruling and the potential difficulty singling out spirits from beer and wine in legislation resulted in no effective effort to reinstate the ban. Though spirits firms were now permitted by DISCUS to advertise on television, national broadcast television networks (NBC, CBS, ABC, and FOX) refused to accept spirits advertising.

Although the first spirits advertisements began appearing on television in 1996, spirits firms spent little on television advertising from 1996-2001. Over this 6 year period, the industry spent \$36 million on television and spent \$ 1.43 billion on magazine advertising, which had dominated their advertising budgets for decades. From 2002-2007 television advertising spending increased 619 percent from \$16 million to \$115 million. Though it is unclear exactly what caused the pause in the adoption of television advertising, fear of public opposition and legislation likely made firms wary to be the first to spend heavily on television. There is also considerable variation in the timing of television advertising adoption across different spirits types and brands. I explore the factors associated with the timing and level of television advertising adoption in this paper.

Framework

General Model of Oligopolistic Advertising Competition

For intuition I follow the general framework of Roberts and Samuelson (1988), which establishes a theoretical model of oligopolistic advertising competition and tests the model on

[2] See Milyo and Waldfogel (1999) for an economic analysis of the ruling's impact on Rhode Island liquor store advertising and prices.

the U.S. cigarette industry. In their model, firms choose a profit maximizing goodwill stock and purchase advertising to maintain that goodwill stock. For simplicity, I assume goodwill stock fully depreciates each period and only consider advertising decisions. Let the profits of firm i in time t be given by

$$\Pi_{it} = p_{it}q_{it} - \psi_i(A_{it}) - c_i(q_{it}), \tag{2.1}$$

where p_{it} is the market price, q_{it} is the quantity demanded from firm i, $\psi_i(\cdot)$ is the advertising cost function for firm i, A_{it} is the level of advertising chosen, and $c_i(\cdot)$ is the firm's production cost function.

Roberts and Samuelson (1988) allow advertising to affect both the size of the market and the shares of the market. Similarly, I define the quantity demanded q_{it} from each firm as the product of the size of the market and the firm's market share:

$$q_{it} = Q_t(AT_t, z_t)S_{it}(A_{it}, AR_t, Z_{it}), \tag{2.2}$$

where $Q_t(\cdot)$ is the total market demand function, AT_t is total advertising by all firms in the market, z_t is a vector of prices and exogenous demand factors unrelated to advertising, $S_{it}(\cdot)$ is the market share function of firm i, AR_{it} is the total advertising of all rival firms, and Z_{it} is a vector of firm-specific exogenous demand factors.

The profit maximizing level of advertising requires that the marginal revenue from advertising equal the marginal cost of advertising:

$$\frac{\partial \Pi_{it}}{\partial A_{it}} = [(p_{it} - \frac{\partial c_{it}}{\partial q_{it}})\frac{\partial q_{it}}{\partial A_{it}}] - \frac{\partial \psi_{it}}{\partial A_{it}} = 0. \tag{2.3}$$

The most interesting element of this equation is the marginal effect of advertising on quantity demanded ($\frac{\partial q_{it}}{\partial A_{it}}$). Using Equation 2, this marginal advertising effect can be decomposed into two elements, one affecting market size and one affecting market share:

$$\frac{\partial q_{it}}{\partial A_{it}} = S_{it}\frac{\partial Q_t}{\partial AT_t} + Q_t\frac{\partial S_{it}}{\partial A_{it}}. \tag{2.4}$$

This equation implies that the return to advertising is a function of the effect of a firm's own advertising on the total size of its market (through total market advertising) multiplied by its market share and the effect on the firm's market share multiplied by the total size of the market. Similarly, the effect of rival firms' advertising on a firm's quantity demanded can be decomposed into two separate effects:

$$\frac{\partial q_{it}}{\partial AR_{it}} = S_{it}\frac{\partial Q_t}{\partial AT_t} + Q_t\frac{\partial S_{it}}{\partial AR_{it}}. \tag{2.5}$$

This implies that rival advertising both helps and hurts a firm. Growing the size of the market helps all firms, but stealing market share hurts rival firms. Whether advertising is cooperative or combative depends on the relative sizes of $\frac{\partial Q_t}{\partial AT_t}$ and $\frac{\partial S_{it}}{\partial AR_{it}}$.

If advertising is cooperative (i.e., it grows market size more so than market share), advertising most benefits firms with large market shares. Firms with small market share may choose not to advertise as their return to growing market size may be smaller than the cost of advertising. This may imply that advertising will be most heavily utilized in highly concentrated markets by firms with a large share of the market. In the case of cooperative advertising, the market goodwill (AT_t) is a public good, which may lead to free-riding and the under-provision of advertising by firms. Assuming there are strictly diminishing returns to total advertising, rival firm advertising will decrease a firm's own advertising.

If advertising is combative and the market share effect dominates the market size effect, then the return to advertising is largely firm specific ($\frac{\partial S_{it}}{\partial A_{it}}$). Additionally, if advertising solely affects market shares, any advertising may be socially inefficient.[3] However, since there are positive returns to advertising for an individual firm, the market may very well end up at an inefficient Nash equilibrium where advertising is over-provided from a social welfare perspective.[4] The collusive spirits broadcast advertising ban was able to partly restrain advertising spending and may have preserved a more efficient outcome than would otherwise have occurred. The lifting of the ban would have disrupted this equilibrium by allowing the use of a new advertising medium. The rate of adoption of television advertising may be determined by the concentration of the market. More concentrated markets, with a few dominant firms or a single market leader, may be able to preserve the higher profit equilibrium without television advertising longer than less concentrated markets.

[3] Advertising may still provide useful information to consumers.

[4] The welfare economics of advertising is controversial and depends on the role of advertising (e.g., persuasive versus informative) and the strategic interactions of firms.

Media Substitution

A key consideration for evaluating the aftermath of the ban is the degree of substitutability across media. If television and magazine advertising are very close substitutes, an advertising ban on one medium will have little effect on sales or advertising spending. If they are poor substitutes, a ban will be a binding constraint on advertising. Seldon and Jung (1993); Seldon, Jewell and OBrien (2000); Färe et al. (2004) find a high level of substitutability among media types. In possibly the only economic study of the end of the spirits broadcast advertising ban, Frank (2008) also finds that print and television advertising are highly substitutable for spirits brands.

Expanding on the previous model, let firms choose two types of advertising, magazine (A_{mit}) and television (A_{vit}). The first order conditions imply that the optimal mix of magazine and television advertising satisfies

$$\frac{\frac{\partial q_{it}}{\partial A_{vit}}}{\frac{\partial q_{it}}{\partial A_{mit}}} = \frac{\frac{\partial \psi_{it}}{\partial A_{vit}}}{\frac{\partial \psi_{it}}{\partial A_{mit}}}. \tag{2.6}$$

This simply states that the ratio of effect on demand of an additional unit of TV advertising to the effect of an additional unit of magazine advertising must be equal to the ratio of the price of a unit of television advertising to the price of magazine advertising. If the return to the first television ad is much greater than the return to the current level of magazine advertising (i.e., $\frac{\partial q_{it}}{\partial A_{vit}}(0) \gg \frac{\partial q_{it}}{\partial A_{mit}}(A_{mit})$), the firm faces a strong incentive to begin advertising on television.

The relative return to television advertising versus magazine advertising may be influenced by a number of factors. Television may be relatively more attractive if consumers (or potential consumers) of a spirits brand watch a large amount of television but read few magazines. Decreasing returns to a single advertising type may also influence a firm's advertising mix. Firms spending heavily on magazine advertising may more readily shift to television to achieve a more optimal mix of advertising. Finally, there are substantial qualitative differences between a static print advertisement and a television ad that allows for movement and sound. Television advertising may be more cost effective for products that can take advantage of the medium. Television advertising may also be more effective for certain groups of consumers, such as youth.

Consumer Demographic and Drinking Characteristics

The advertising decisions of firms may depend on characteristics of the consumers of that brand or the consumers the firm seeks to attract. For example, the drinking preferences of a brands consumers may affect whether it views advertising as combative or cooperative. If consumers of a particular type or brand of spirits also consume other alcohol types (i.e., beer and wine), they may be convinced by advertising to choose a bundle of alcohol that contains a greater proportion of spirits. On the other hand, spirits brands whose current consumers seldom drink alcohol of other types may want to seek out consumers who are less aware of the merits of that particular type and who may receive more information from advertising for spirits of that type. Consequently, I cannot strongly predict whether advertisers will target those who drink a greater percentage of their alcohol as spirits or those who drink a smaller percentage.

Theory also provides no strong predictions regarding the relationship between other demographic characteristics of the consumers of an alcohol brand and that brand's advertising choices. As previously discussed, television advertising may be more effective on youth relative to advertising in magazines. Additionally, advertising may affect youth more in general if they are less informed about alcohol brands or experience a greater complementarity between advertising and consumption. However, brands more popular with youth are likely sensitive to the scrutiny their advertising is under from government and public health advocates. The expected outcry if alcohol firms are perceived as marketing to underage youth may constitute an additional non-monetary cost of advertising. Given the controversy over ending the television advertising ban, this cost may be higher for television advertising than magazine advertising and may cause brands popular with youth to avoid advertising on television. As these two proposed effects operate in opposite directions, I make no prediction of whether brands popular with youth are more or less likely to advertise on television.

Similar to the examples of the drinking preferences and age of a brand's consumers, it is difficult to predict whether and how advertising decisions may be related to several other consumer demographic factors, such as gender, race, education, income, and student status. I include variables measuring these factors in my models for three reasons. First, they likely are correlated with drinking behavior and media viewing, and omitted the additional demographic factors may bias my estimates of the relationship between advertising and other variables. Second, there is concern that alcohol firms target advertising at some of these groups, particularly students and African-Americans (Kuo et al., 2003; Hackbarth, Silvestri and Cosper, 1995). Finally, the advertising choice functions of alcohol firms are unobserved, making it an unresolved question whether these factors are related to advertising decisions.

2.3 Data

Kantar Media Magazine and Television Advertising Data

I use data measuring advertising spending on television from 1995 to 2010 compiled by Kantar Media.[5] For each television advertisement airing, the television data include: the alcohol brand, the sponsor name, the name of the program during which it aired, when it aired (date and time), the coverage of the message (national broadcast TV, local broadcast TV, or cable TV), its length (in seconds), the geographic media market where it aired, and, most importantly, the estimated cost of airing the ad. The television data track messages aired nationally or aired locally in 110 DMAs. The data represent the universe of television alcohol advertisements appearing nationally on cable and broadcast network and locally on broadcast networks.[6] I use these data to construct a series measuring yearly brand-level spending on television advertising.

I also use data from Kantar Media measuring advertising spending in magazines from 1995 to 2010. For each magazine advertisement the data include: the alcohol brand and sponsor, the magazine, the issue (date), the page size of the advertisement, and the estimated cost of placing the ad. I use these data to construct an analogous series measuring yearly brand-level spending on magazine advertising. Figure 2.1 shows the general trends in yearly spending on spirits advertisements in magazines and on television. The data show a leveling off then decline in magazine advertising and a sharp increase in television advertising. This increase in television advertising is particularly sharp in 2002 and subsequent years.

Simmons National Consumer Survey (NCS)

I use data from the Simmons National Consumer survey from 2000-2007 to generate the demographic, media viewing, and drinking characteristics of consumers of each spirits brand. The NCS is a biannual survey of consumer behavior. Simmons uses stratified probability sampling to draw a sample that is representative of the population of the contiguous United States. The NCS collects detail information on the consumption of a nearly exhaustive list

[5]Kantar Media was previously known as TNS Media and Competitive Media Reporting (CMR).

[6]Kantar does not measure local cable advertising and only measures cable advertising on 44 popular, national cable networks.

of products, demographic characteristics, and media (television, radio, print, and internet) viewing. Specifically, the NCS asks respondents to report the types and brands of alcohol they consume and how much of each type and brand they consumed in the past 30 days. I use these data to construct the demographic characteristics (gender, age, race, education, income, and student status) and media viewing behavior (number of issues of magazines read, hours of cable television watched per week, and hours of network television watched) of each brand of spirits listed in the survey. I also generate summary statistics of the drinking behavior of the consumers of each brand: the average number of beers consumed in the last 30 days, the average drinks of wine, and the average drinks of spirits.

Adams Liquor Handbook

The Adams Liquor Handbook, published by the Adams Beverage Group, includes detailed data measuring numerous aspects of the spirits industry (Adams Beverage Group, 2006*b*). Data measuring the number of cases sold of every major spirits brand allow me to construct the total yearly sales of each spirits type (in 9-liter cases) and the yearly market shares of each brand.

I use the market shares from these data to generate type the one-firm concentration ratio for each spirits (i.e., the market share of the brand with the highest market share), the two through four-firm ratio (i.e., the total market share of the brands with the second, third, and fourth highest market shares), and an indicator of whether a brand has the highest market share of that type (the market leader). These measures of market share have notable drawbacks compared to the Herfindahl Hirschman Index (HHI). Specifically, the first and second through third firms are not particularly special and may give only a partial summary of a market's concentration. The HHI takes all firms in the market into account and provides a more complete measure of market concentration. However, I use concentration ratios for two reasons. First, I have market share data only for the most popular brands in the market, which would require me to make assumptions about the market shares of less popular brands to approximate the HHI. Second, splitting the measures of market share into the one-firm and two through four-firm ratios allows me to test for separate effects of market concentration, as summarized by these two measures, on advertising. That is, I can test if the market share of the top firm has a different relationship to advertising than the market share of firms two through four. Using the HHI only allows one to test for a general relationship between market concentration and advertising.

I merge these yearly brand-level consumer and market characteristics with the Kantar advertising series to create a brand-level series of advertising spending and brand and market characteristics. I also generate several discrete outcomes measuring television advertising adoption. I create a variable indicating the first year a brand "started" advertising on television, defined as the first year the brand spent more than $500,000 on television advertising. I generate an indicator of the first year a brand "switched" to television advertising, defined as the first year the brand spent more on television advertising than in magazines. Finally, I define a brand as a "first mover" if it was the first brand of its type to "start" television advertising.

2.4 Methods and Identification

Comparison of Means

I first compare the means of consumer and market characteristics of brands that ultimately began television advertising and those that do not. To avoid the simultaneity of advertising affecting consumption and consumer characteristics affecting advertising strategies, I compare brand and market characteristics in 2000, before the significant adoption of television advertising, of brands that eventually "started" with those that did not. In other tables, I also compare the year 2000 characteristics of brand that eventually "switched" to television advertising to non-switchers and compare "first-mover" brands to other brands.

Discrete Models of Advertising Entry

As described earlier, I define three discrete measures of television advertising adoption: "starting", "switching", and being the "first mover" of a spirits type . I first model whether a brand ever starts, switches, or moves first over the 2001-2007 time period as a function of the year 2000 characteristics of the brand and its market. I use the general specification

$$Y_i = \beta_0 + demo_{i,t=2000}\beta_1 + drink_{i,t=2000}\beta_2 + mkt_{i,t=2000}\beta_3 + \epsilon_i, \tag{2.7}$$

where Y_i is starting, switching, or first-moving to television advertising; $demo_{i,t=2000}$ refers to the year 2000 demographic and media viewing characteristics (age, gender, magazine reading, and television viewing); $drink_{i,t=2000}$ refers to the alcohol consumption characteristics of consumers of that brand (age, gender, race, education, household income, and student status); and $mkt_{i,t=2000}$ refers to the year 2000 market characteristics of the brand's type.

Though isolating a causal effect of advertising is very challenging, my approach attempts to address the largest threat to threat to causal inference, the simultaneity of advertising and sales. Firms make advertising decisions based on their sales and market share, and those advertising decision may subsequently affect their sales and market share. Though one clearly follows the other temporally, identifying an advertising effect requires strong assumptions about the timing of the effect and the precise knowledge of when advertising is placed and when sales occur. Because I use the initial (Year 2000) brand and market characteristics, I avoid the potential for reverse causality–post ban advertising cannot possibly affect market shares during the ban. Though I am cautious in interpreting my results as causal, I believe that this approach may shed some light on whether market structure affects incentives to advertise.

Panel Models of Advertising Spending

Next, I estimate panel models of television advertising spending. I include lagged vectors of brand demographic and media viewing characteristics, drinking behavior, and market characteristics that vary by brand (i) and year (t). I use lagged values under the assumption that firms can only use last year's brand and market characteristics to plan this year's advertising. I also interact these vectors with the year to allow the relationship between advertising and these characteristics to change linearly over time. I use the general specification

$$\begin{aligned} Y_{it} = \gamma_0 &+ demo_{it-1}\gamma_1 + demo_{it-1} \cdot t_t\gamma_2 + drink_{it-1}\gamma_3 + drink_{it-1} \cdot t_t\gamma_4 \\ &+ mkt_{it-1}\gamma_5 + mkt_{it-1} \cdot t_t\gamma_6 + spend_{it}\gamma_7 + \epsilon_i, \end{aligned} \tag{2.8}$$

where Y_{it} is either spending on television advertising (in \$ millions) or the fraction of advertising spending used on television and $spend_{it}$ is a vector of current and lagged television advertising spending and adoption by market competitors. These models explore whether the relationship between brand-level television advertising and brand and market characteristics changes over time. That is, were the first brand to advertise spirits on television different from those who advertise later?

Hazard Models of Advertising Entry

Finally, I estimate hazard models of "starting" and "switching" to television advertising. These differ from the first set of models since they dynamically model the choice to begin television advertising. The relationship between the adoption of television advertising by rival brand and the decision to "start" or "switch" is of particular interest. Formally, I estimate a discrete-time hazard model of these events as linear function of the same set of variables used in the panel models of advertising spending:

$$\begin{aligned} Y_{it} = \gamma_0 &+ demo_{it-1}\gamma_1 + demo_{it-1} \cdot t_t\gamma_2 + drink_{it-1}\gamma_3 + drink_{it-1} \cdot t_t\gamma_4 \\ &+ mkt_{it-1}\gamma_5 + mkt_{it-1} \cdot t_t\gamma_6 + spend_{it}\gamma_7 + \epsilon_i, \end{aligned} \tag{2.9}$$

where Y_{it} is a discrete variable measuring either "starting" or "switching to" television advertising as defined earlier.

2.5 Results

Figure 2.1 and Figure 2.2 show general trends of magazine and television spirits advertising and the percentage of advertising spending on television. There is significant growth in advertising spending on television from 2001 to 2006. Figures 2.3 and 2.4 show the total trends by spirits type for "brown" spirits. These types show modest increase of most of the 2000's with several types spending close to 50% of their budget on television by the end of the decades. Figures 2.5 and 2.6 give trends for "clear" spirits and cordials[7]. These types, particularly rum and vodka, show high levels of spending on television advertising, which is perhaps not surprising given the high levels of total advertising in these types.

Figures 2.7 to 2.14 show the amount of television advertising spending and the percentage of advertising on television for brands with more than 5% market share of their spirits type. These figures show large differences between brands in the same market in the timing and degree of television advertising adoption.

Table 2.1 compares the mean year 2000 characteristics of brands that chose to "start" television advertising at some point with those brands that do not start. I find that consumers of brands that started TV advertising were significantly better educated and consume less alcohol of all types. They also read fewer magazine issues and watch less cable television.

[7]Cordials include liqueurs and drink mixers with lower alcohol content.

These brands also have higher market share and, subsequently, higher sales. They are much more likely to have been the market leader and advertised much more heavily in magazines in 1996. The differences between brands that switched to television and those that did not are largely similar (Table 2.2). Switchers have, on average, more educated consumers that drink more moderately and have higher sales and market share. Table 2.3 compares brands that were the first to begin advertising on television to those that started later or not at all. These first-movers have, on average, consumers that are less likely to read magazines and watch television.

Table 2.4 lists the number of brands analyzed in each spirits type, the percentage of those brands that ever started and switched to television advertising, and the brand(s) that was the first mover in the category. There is a large degree of variation in the adoption of television advertising across different spirits types. Though only 25% of bourbon brands (two of eleven) start advertising on television, 50% (three of six) rum brands begin advertising on television. There is a similar variance in the percentage of brands switching to television advertising, with half of rum brands switching but no cognac or blended whiskey brands making the switch. Although a large majority of the brands that start advertising on television eventually switch to television advertising, the timing of these changes differs considerably and makes studying both transitions useful.

Table 2.5 presents regression coefficients modeling whether a spirits brand eventually started, switched to, or was the first of its type to start advertising alcohol on television.[8] I model the relationship between these outcomes and the year 2000 brand and market characteristics. There is a strong relationship between measures of market competition and adoption. Market leading firms are less likely to start advertising when they have a high market share and face little competition. Several of the coefficients measuring the relationship between market competition and television advertising are difficult to interpret on their own, due to multiple interactions. Because the one-firm concentration ratio is the market share of the market leader, the proper relationship between an increase in the one-firm ratio and the dependent variable for the market leader is the sum of the coefficients of the one-firm concentration ratio, the market share, and the interaction of the market leader indicator and the one-firm ratio. This implies that a 10 percent increase in the one-firm ratio is associated with a 0.02 percentage point decrease in the likelihood a market leader will start television advertising.[9] However, this relationship is not significantly different from zero ($p = 0.998$).

[8] Appendix A includes alternative specifications that add several interaction terms sequentially and estimate separate models by market rank categories.

[9] This holds constant the two to four-firm ratio, assuming that the increase in market share comes at the

Similarly, the relationship between the two to four-firm ratio and starting television advertising is the sum of the two to four-firm ratio and its interaction with the market leader indicator. This implies a 10 percentage point increase in the two to four-firm ratio is associated with a 53 percentage point increase in the likelihood a brand starts TV advertising ($p < 0.01$) and a 62 percentage point increase in the likelihood of switching to TV ($p < 0.01$). The main coefficient on the one-firm concentration ratio gives the relationship between the ratio and the dependent variable for brands that are not the market leader. It implies a 10 percentage point increase in the one-firm ratio is associated with a 3.6 percentage point increase in the likelihood of starting alcohol advertising.

The results imply that few consumer demographic and drinking factors have a statistically significant relationship with the adoption of TV advertising. There is no significant relationship between the percentage of underage or young overage drinkers of a brand and television advertising. Brands more popular with beer and wine drinkers are less likely to start or switch. However, only the beer result is statistically significant. There is also no significant relationship between media viewing characteristics and TV advertising adoption.

In tables 2.6, 2.7, and 2.8, I simulate the relationship between various one and two-to-four firm ratios and the change in the predicted probability a brand ever advertises on television. I use the results from Table 2.5 to predict the probability a brand ever advertises on television given various combinations of one and two-to-four firm ratios. I then average the predicted probabilities for every brand in three different market rank categories: market leader, firms ranked two through four, and firms ranked five or lower. Table 2.6 shows that the probability a market leading firm with a 20 percent market share and a 20 percent two-to-four firm ratio has a 30 percent probability of ever using television advertising. This probability weakly declines as the leading firm's market share increases (moving across each row). However, there is a large increase in the probability a market leading firm ever advertises as the two-to-four firm market share increases. A market leading firm in a market with a two-to-four firm ratio of 40 percent has a probability of ever advertising on television greater than one, and firms ranked lower than four have negative probabilities of advertising. These predicted probabilities greater than one and less than zero are due to the use of linear probability models and reflect very high and very low probabilities of advertising. There are three markets with a greater than 40 percent two-to-four brand share: Canadian whisky, bourbon whiskey, and Irish whiskey. These markets are generally small in terms of sales and are comprised of a few brands that dominate the market. Table 2.7 shows a strong

cost of brands outside of the top four.

negative relationship between the two-to-four firm ratio and television advertising for brands ranked two through four, holding a brand's own market share constant. Table 2.8 presents simulations for firms with small market shares. Brands facing large market leaders are more likely to advertise, but higher two-to-four brand market shares make it much less likely that these brands will advertise on television.

Table 2.9 presents the relationship between lagged yearly brand and market characteristics and yearly spending on television advertising. In 2001, brands that lead their market spend less on television advertising. However, the higher their own market share or the share held by the second through fourth most popular brands of its type the more heavily the leading brand spends on television advertising. Initially, there is a relationship between market share and advertising spending for the market leading brand. A 10 percent increase in market share is associated with a \$ 0.12 million increase in yearly advertising spending for a market leader. A 10 percent increase in the share of the market held by the second, third, and fourth brands is associated with a \$ 0.19 million increase in advertising spending in 2001. These relationships both subsequently decrease over time. There is also a strong serial correlation in television advertising spending. On average, brands spend 49 percent more than in the previous year, though these increases in spending diminish over time. There is also a significant relationship between a brand's television advertising and the television advertising of its rival brands in the previous years. On average, a brand matched 14 percent of the total advertising spent by its rivals. Initially, brands that are more popular with underage drinkers and 21- to 30-year olds spend less heavily on television advertising. However, as time progresses brands popular with these age groups spend more heavily on television advertising.

Table 2.10 presents coefficient estimates of hazard models of starting and switching to television advertising. Brands with higher market shares are also less likely to start advertising earlier. However, holding market share constant, brands with higher sales are more likely to start or switch earlier. There is a statistically weak relationship with rival brand television advertising in the previous year. A one million dollar increase in rival television advertising is associated with a 2.0 percentage point increase in the likelihood a brand starts advertising on television in given year. The results imply that brands that are popular with younger drinkers are no more likely to start advertising on television earlier. A 10 percentage point increase in the fraction of drinkers of a brand that are underage is associated with a 0.6 percentage point increase in the likelihood of starting television advertising in a given year. Brands popular with liquor drinkers are less likely to start advertising earlier but are more likely to start by 2005. Brands popular with beer and wine drinkers are more likely to start advertising initially.

2.6 Discussion

While the results of this study are not meant to estimate the causal effect of brand and market characteristics on brand-level television advertising by spirits firms, the estimates do provide some insight into the relationship between these factors and advertising. Though theory may predict that certain market factors and brand characteristics should be related to advertising decisions, this study empirically tests whether they are related in a very specific context, the adoption of television advertising for spirits. Additionally, a number of my analyses study the relationship between market factors prior to the lifting of the ban and post-ban advertising behavior, eliminating the largest threat to a causal interpretation of my results, simultaneity. Market characteristics, including the level of competition in a market, appear to play a significant role in the degree and timing of the adoption of television advertising. However, the results imply that the characteristics of a brand's consumers are only weakly related to television advertising adoption. In addition, little seems to predict the timing of the adoption of television advertising, which may largely be determined by firm idiosyncrasies.

The strongest correlates with television advertising adoption are the competitiveness of a market and the advertising of rival brands. The results suggest that advertising may play an important competitive role in spirits markets. Smaller brands are more likely to advertise when a market leader holds a large portion of the market share and less likely to advertise when the second, third, and fourth most popular brands hold a larger share. This perhaps implies a larger role for television advertising in markets that have one dominant brand and a number of smaller, but similarly-sized brands. Market leading brands are more likely to adopt television advertising when they face greater competitive pressure. The larger the share of the market leading brand, the less likely that brand is to advertise. However, holding its own market share constant, the higher the share of the next three most popular brands, the more likely the market leader is to use television advertising. This result is strongly statistically significant and consistent with spirits advertising being more combative than cooperative. That is, it has a larger effect on market share than market size. If firms anticipated that television advertising has a strong effect on market size, the market leading firm with large market share, who are able to capture most of the returns to growing a market's size, should be *more* likely to advertise. There is also a positive relationship between rival brand advertising in the previous year and a brand's own advertising. Brands spend more on television advertising when their competitors spent more last year. Again, this result is consistent with a dominating combative effect of television advertising. If advertising were largely cooperative (and therefore a public good), firms might potentially free-ride off the advertising goodwill created by their rival's marketing in the previous year.

There is less evidence of a relationship between the demographic characteristics and alcohol consumption of a brand's consumers and the brand's adoption of television advertising. The results suggest that the brands that adopted television advertising earlier were more popular with 18- to 20-year-olds and 21- to 30-year-olds. However, these results are not statistically significant, and these brands are less likely to move to television overall. This may calm concern in the public health community, which protested the end of the television advertising ban due to concern that firms would market spirits to the underage. I do find some evidence that the education level of a brand's consumers is associated with television adoption. Brands with higher levels of consumers with college degrees are more likely to start and switch to television advertising. Based on the theory outlined earlier, it is unclear why brand with more educated consumers would find higher returns to advertising on television. Perhaps these consumers are more highly sought after by firms due to their higher spending power, or maybe they are more swayed by television advertising (or advertising in general).

I find some evidence that the drinking behavior of the consumers of a brand influences the adoption of television advertising. Spirits firms cited the loss of market share to beer and wine and a desire for a "level playing field" as reasons for lifting the ban. This suggests that brands that see themselves as stronger substitutes for beer and wine might choose to advertise on television in an attempt to steal back market share from these products. While not statistically significant, I find evidence that brands whose consumers drink more beer and wine are *less* likely to ever start or switch to television advertising, and brands whose consumers drink more liquor are more likely to ever substantially adopt TV advertising. Though again, only the beer result is statistically significant. However, I do find some evidence that brands less popular with spirits drinkers are significantly more likely to start advertising earlier. Overall, these results do not suggest that competition with beer and wine producers was a strong motivator for the adoption of television advertising by spirits firms.

The media viewing characteristics of consumers also are not strongly related to television advertising adoption. There is no significant relationship between the magazine reading and television viewing of consumers and a brand starting or switching to television advertising. The mean comparisons and regression results in table 2.5 imply that brands with consumers that read fewer magazine issues are more likely to start and switch to magazine advertising. However, the consumers of these brands also watch less television. Consumers of these brands may be simply more difficult to reach by advertising in general since they spend less time reading and watching TV, but it may be more cost-effective to reach them with television advertising than magazine advertising.

Though this study provides a valuable, sophisticated description of advertising following the lifting of the spirits broadcast advertising ban, it has significant limitations. As previously mentioned, I intend this study to be an in-depth description of which spirits brands took up television advertising and the factors that correlate with this adoption. I do not claim that these correlations imply a causal relationship. Attempting to estimate causal pathways would require imposing significant structure on a model of advertising and additional data, notably price data. Structural modeling could address the problem of simultaneity in the panel and hazard models. If advertising affects market shares and market shares affect advertising, then the direction of causality cannot be determined in simple OLS models. Determining the causal effect of market concentration on advertising would require an instrument for market concentration. In the specific context of the spirits market, there is not an obvious instrument that affects market shares but has no additional effect on advertising. Simultaneity is less of an issue with the models I estimate using year 2000 market characteristics. However, these models may still suffer from omitted variables bias. There may be unobserved characteristics correlated with both brand and market characteristics and advertising. The largest challenge to causal inference is the limited number of markets (spirits types) and brands in those markets. I cannot estimate models with type fixed effects because I only observe each spirits market experiencing the lifting of the broadcast ban once. It could be that the taste and experience of rum lends itself to more concentrated markets and television advertising. However, there is no accounting for (or controlling for) tastes, and I observe no counterfactual, more diffuse rum market. No structural model or additional data can solve this problem. Comparisons to markets for other goods can provide additional data, but the unique circumstances of the lifting of the spirits broadcast ban may limit the usefulness of such comparisons. The lack of statistically significant correlates with television advertising adoption, particularly the timing of adoption, suggests adoption may be largely due to market and brand idiosyncrasies that cannot be accounted for with my data. Despite these limitations my analyses provide important preliminary insight into the ban's aftermath.

The results do not seem to contradict the notion that the collusive ban largely kept spirits brands from competing among themselves. Market leading brands, which had the most to lose from declining liquor sales, do not seem to upset the status quo and adopt television advertising unless competition forces their hand. Though spirits advertising may be more combative than cooperative, total sales of spirits have increased in the past decade along with the advertising budgets of spirits firms. However, it is beyond the scope of this study to conclude the increase in spirits consumption is the result of increased advertising.

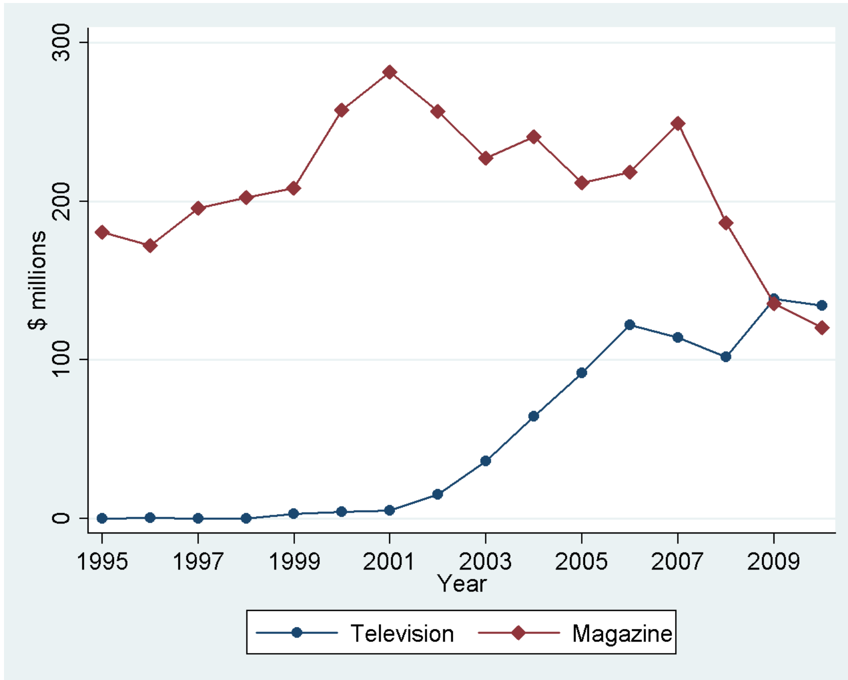

Figure 2.1: Total spending on spirits advertising, in millions of dollars, in all magazines and on all television programs tracked by Kantar Media, 1995-2010.

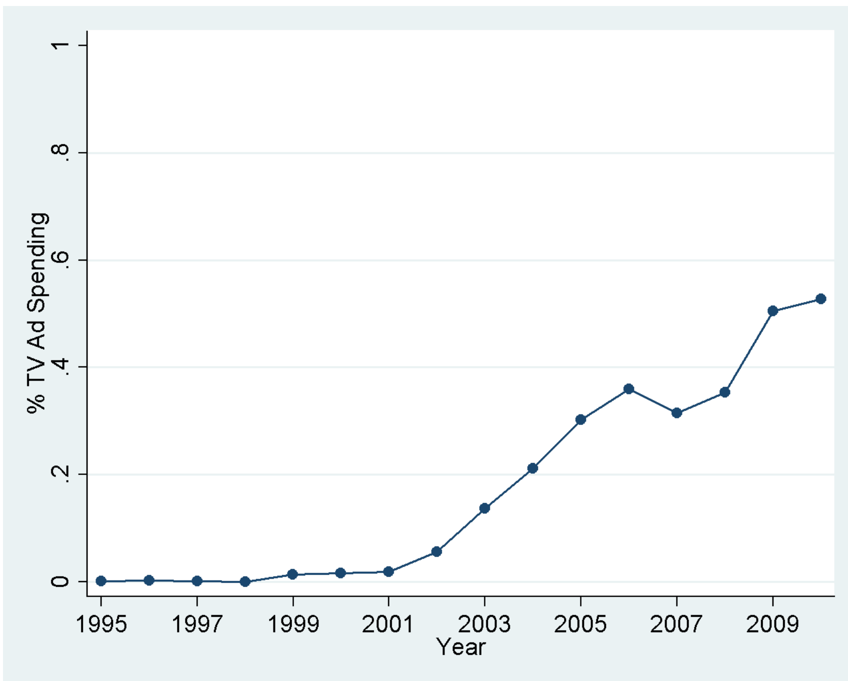

Figure 2.2: Total percentage of television and magazine advertising spending allocated to television advertising. Source: Kantar Media 1995-2010.

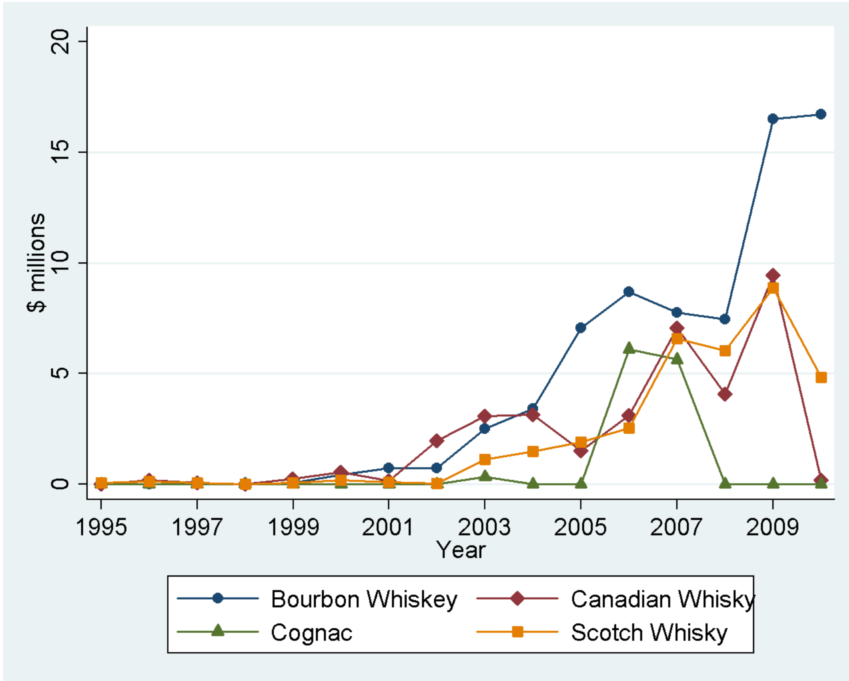

Figure 2.3: Spending on "brown" spirits advertising, in millions of dollars, in all magazines and on all television programs tracked by Kantar Media, 1995-2010.

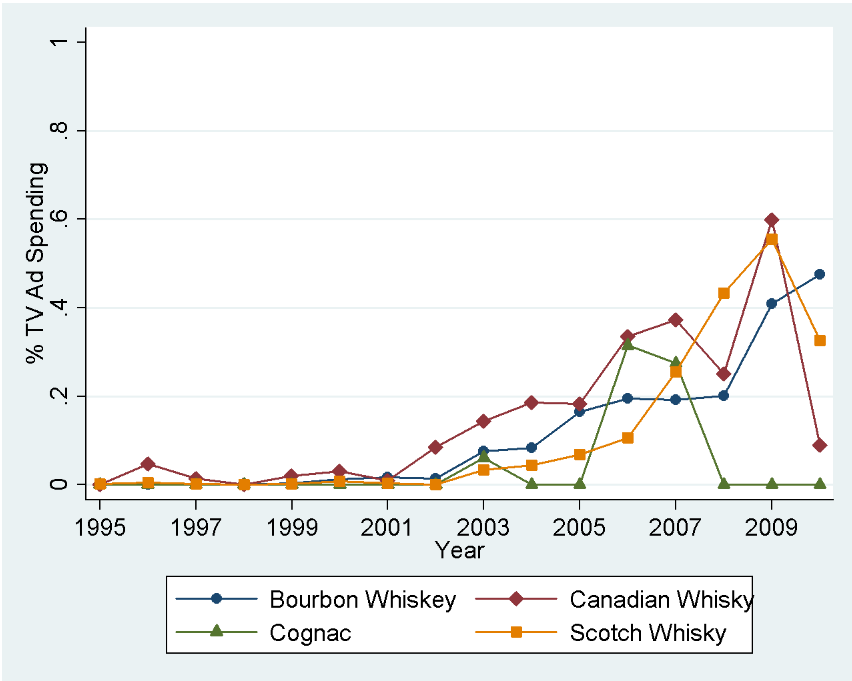

Figure 2.4: Percentage of "brown" spirits television and magazine advertising spending allocated to television advertising. Source: Kantar Media 1995-2010.

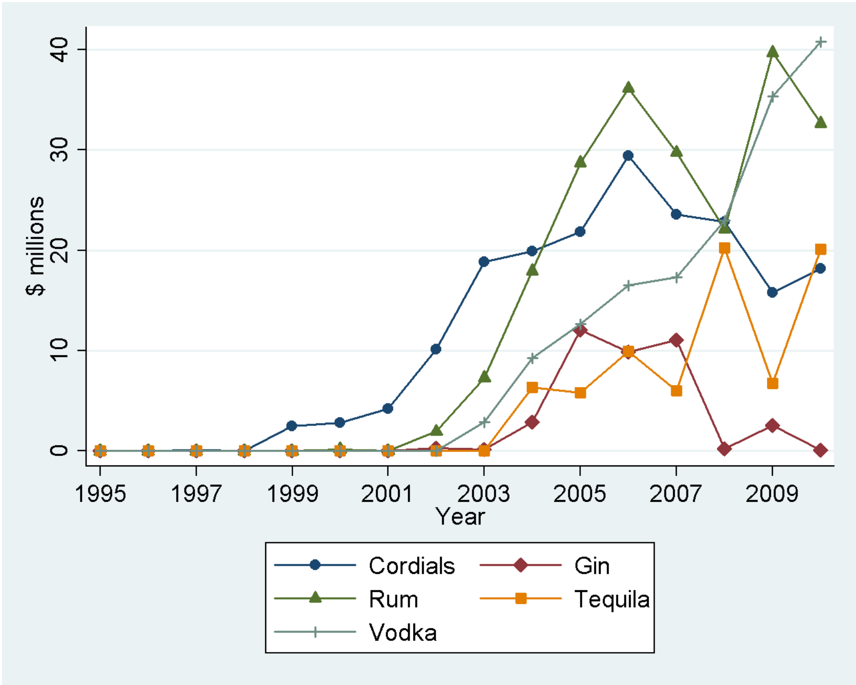

Figure 2.5: Spending on "clear" spirits advertising, in millions of dollars, in all magazines and on all television programs tracked by Kantar Media, 1995-2010.

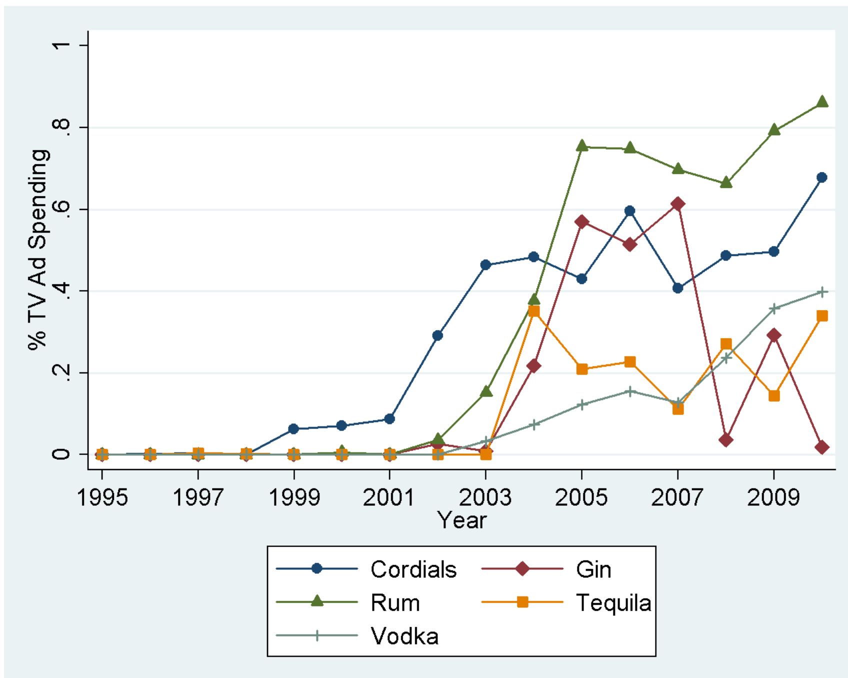

Figure 2.6: Percentage of "clear" spirits television and magazine advertising spending allocated to television advertising. Source: Kantar Media 1995-2010.

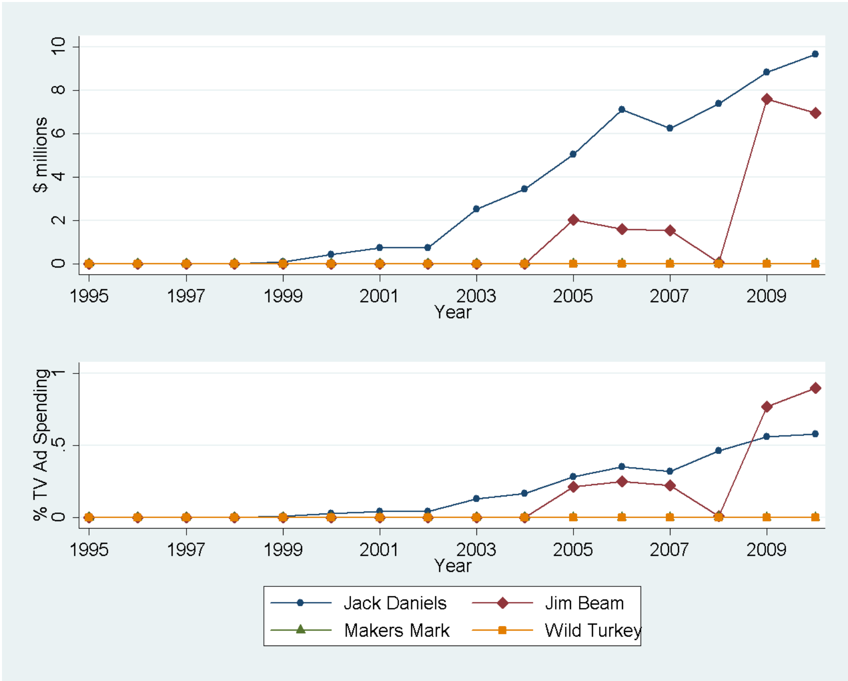

Figure 2.7: Spending on bourbon whiskey television advertising by brand, in millions of dollars and as a fraction of television and magazine advertising spending. Source: Kantar Media, 1995-2010.

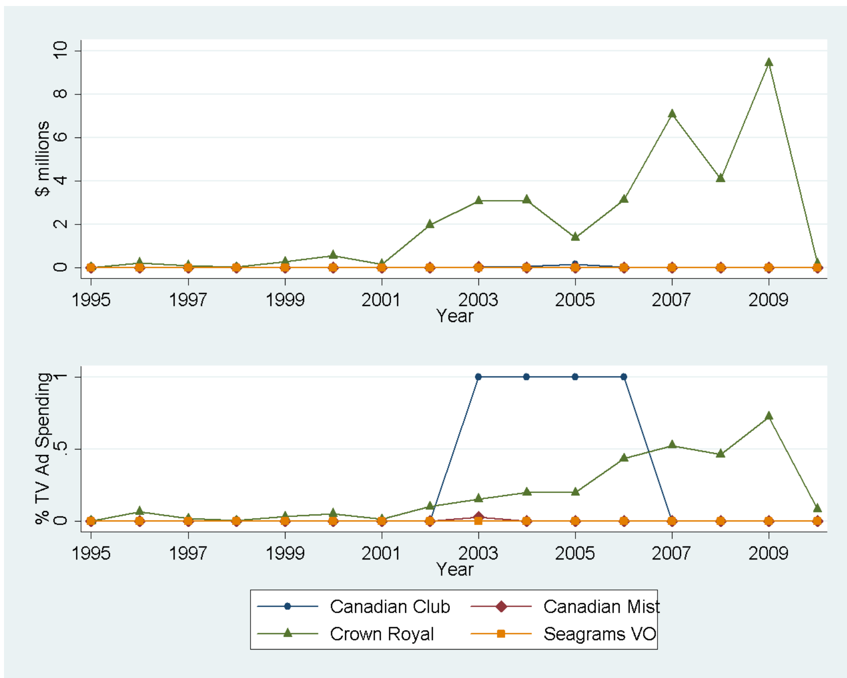

Figure 2.8: Spending on Canadian whisky television advertising by brand, in millions of dollars and as a fraction of television and magazine advertising spending. Source: Kantar Media, 1995-2010.

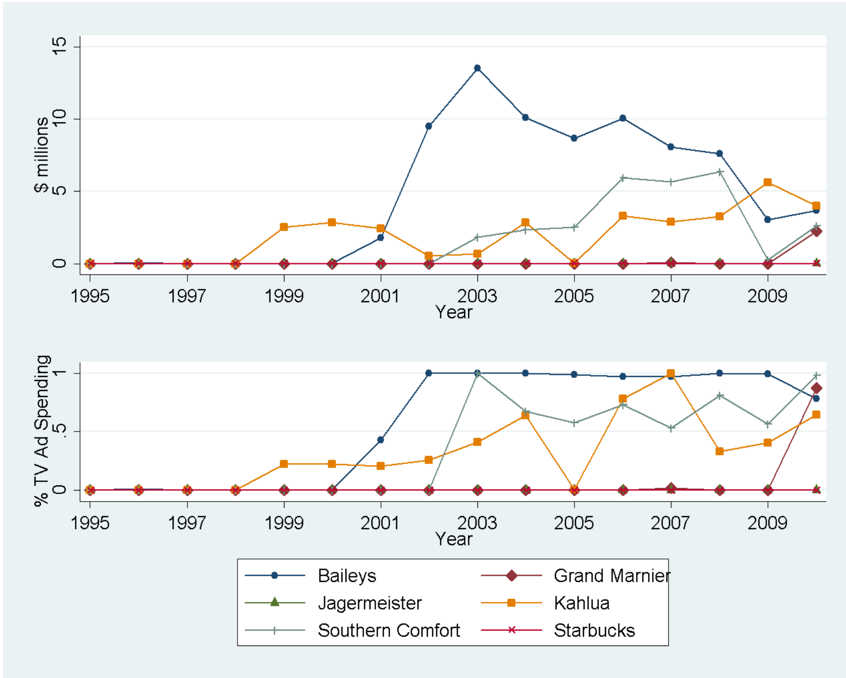

Figure 2.9: Spending on cordial television advertising by brand, in millions of dollars and as a fraction of television and magazine advertising spending. Source: Kantar Media, 1995-2010.

Figure 2.10: Spending on gin television advertising by brand, in millions of dollars and as a fraction of television and magazine advertising spending. Source: Kantar Media, 1995-2010.

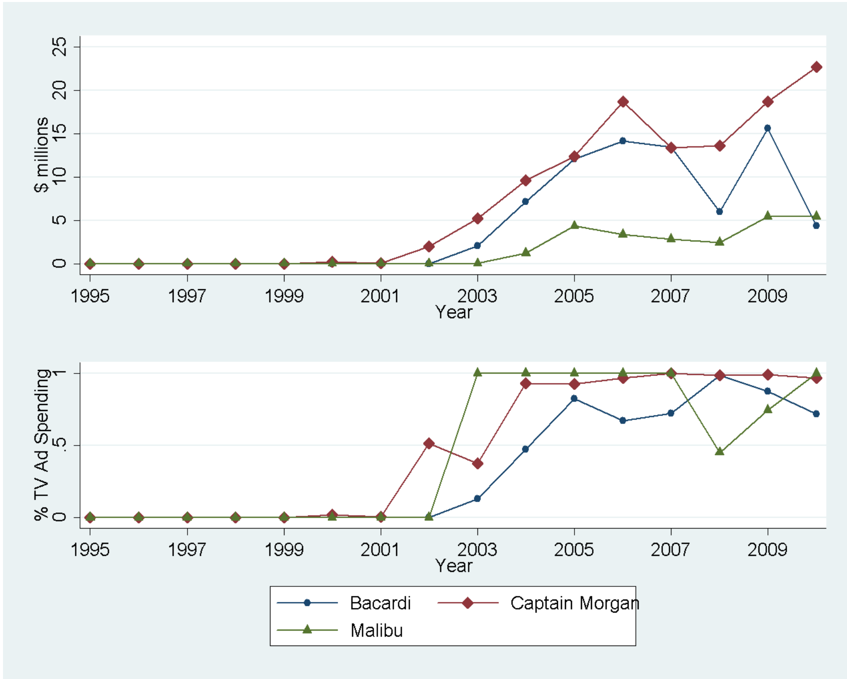

Figure 2.11: Spending on rum television advertising by brand, in millions of dollars and as a fraction of television and magazine advertising spending. Source: Kantar Media, 1995-2010.

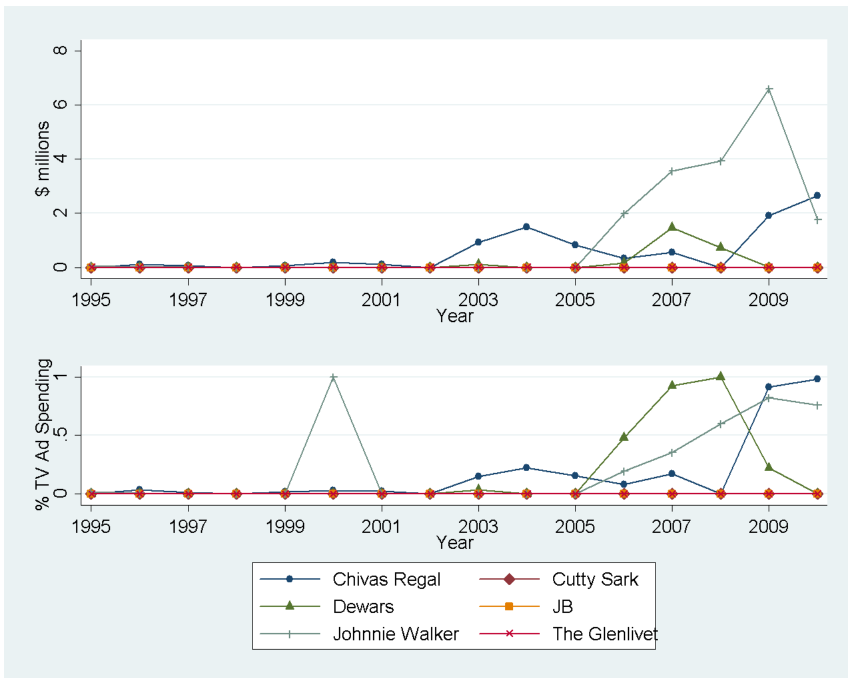

Figure 2.12: Spending on Scotch whisky television advertising by brand, in millions of dollars and as a fraction of television and magazine advertising spending. Source: Kantar Media, 1995-2010.

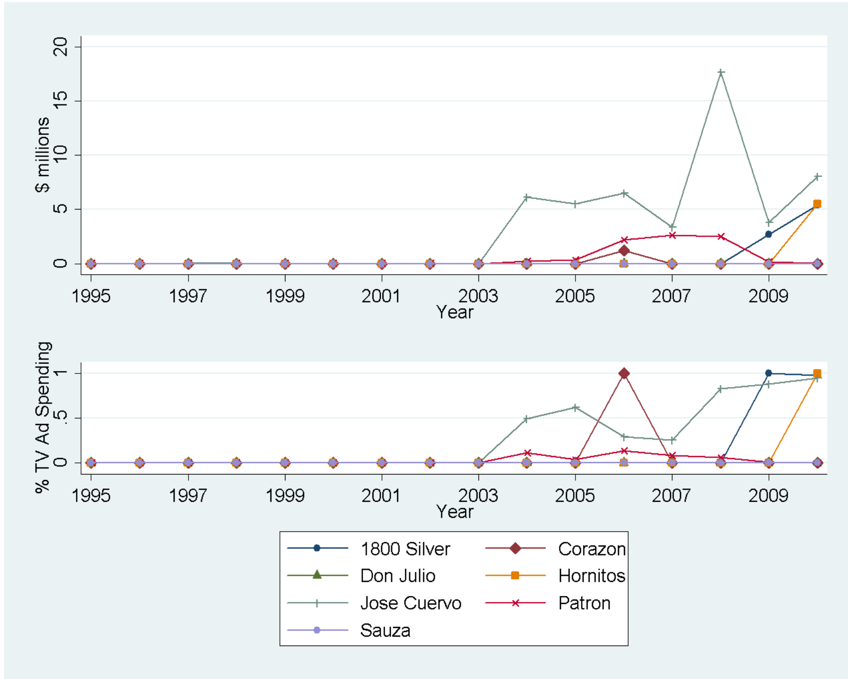

Figure 2.13: Spending on tequila television advertising by brand, in millions of dollars and as a fraction of television and magazine advertising spending. Source: Kantar Media, 1995-2010.

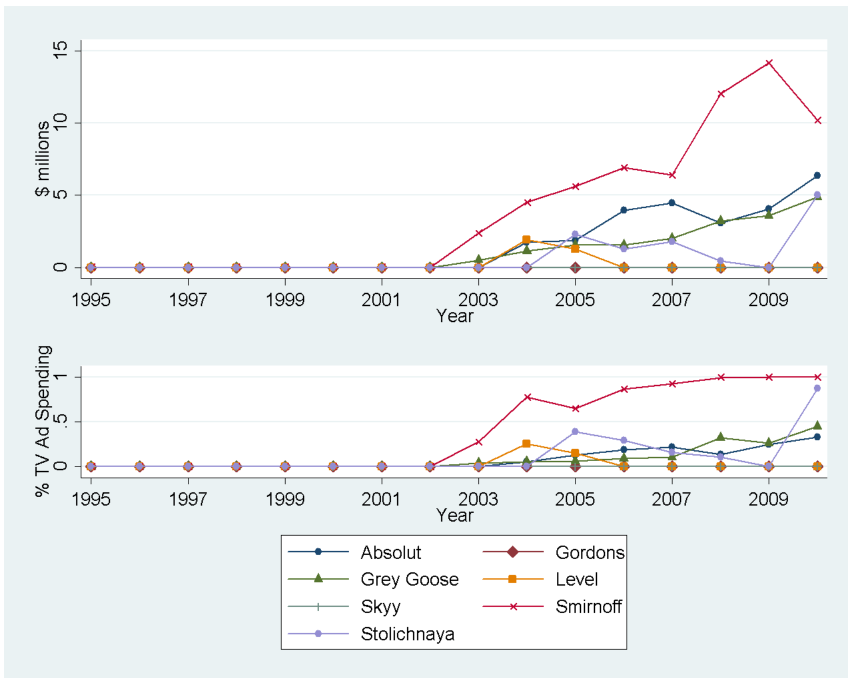

Figure 2.14: Spending on vodka television advertising by brand, in millions of dollars and as a fraction of television and magazine advertising spending. Source: Kantar Media, 1995-2010.

Table 2.1: Comparision of Starters to Non-Starters

	Non-Starters	Starters	Diff. T-Stat
Pct. Female	0.40	0.41	-0.45
Pct. Age 18-20	0.06	0.04	1.32
Pct. Age 21-30	0.18	0.23	-1.84
Pct. Black	0.12	0.10	1.33
Pct. Other	0.03	0.03	0.08
Pct. Hispanic	0.16	0.14	1.20
Pct. Less than HS Educ.	0.18	0.11	4.49***
Pct. Some College Educ.	0.27	0.30	-2.43*
Pct. College Grad. Educ.	0.26	0.31	-2.66**
Avg. HH Income	81.39	89.39	-2.70**
Pct. Student	0.09	0.09	-0.35
Avg. Liquor Drinks	39.64	28.00	4.32***
Avg. Beer Drinks	22.94	19.02	3.22**
Avg. Wine Drinks	14.37	10.66	3.62***
Avg. Magazine Issues	271.44	205.40	3.14**
Avg. Hours TV per Week	57.10	47.90	3.20**
One Firm Concentration Ratio	0.23	0.23	0.01
Two to Four Firm Ratio	0.28	0.26	0.80
Mkt. Share in Type	0.06	0.12	-2.52*
Sales of Brand (000 Cases)	0.61	1.71	-3.43**
Total Sales of Type (000 Cases)	198.71	210.55	-0.67
Market Leader	0.07	0.25	-2.22*
Magazine Spending in 1996	0.46	5.09	-5.02***
Observations	72		

Table 2.2: Comparision of Switchers to Non-Switchers

	Non-Switchers	Switchers	Diff. T-Stat
Pct. Female	0.40	0.42	-0.97
Pct. Age 18-20	0.06	0.04	1.42
Pct. Age 21-30	0.19	0.22	-1.32
Pct. Black	0.11	0.10	0.85
Pct. Other	0.03	0.03	-0.35
Pct. Hispanic	0.16	0.13	1.78
Pct. Less than HS Educ.	0.17	0.11	4.04***
Pct. Some College Educ.	0.27	0.29	-1.75
Pct. College Grad. Educ.	0.26	0.32	-2.64*
Avg. HH Income	82.24	88.75	-2.11*
Pct. Student	0.09	0.09	0.23
Avg. Liquor Drinks	38.84	28.09	3.80***
Avg. Beer Drinks	22.83	18.76	3.27**
Avg. Wine Drinks	14.11	10.71	3.18**
Avg. Magazine Issues	266.63	206.52	2.75**
Avg. Hours TV per Week	56.12	48.64	2.48*
One Firm Concentration Ratio	0.24	0.23	0.19
Two to Four Firm Ratio	0.27	0.27	0.46
Mkt. Share in Type	0.06	0.13	-2.83**
Sales of Brand (000 Cases)	0.68	1.71	-3.12**
Total Sales of Type (000 Cases)	202.06	205.66	-0.20
Market Leader	0.06	0.28	-2.61*
Magazine Spending in 1996	0.94	4.73	-3.77***
Observations	72		

Table 2.3: Comparision of First Movers to Non-First Movers

	Not First	First	Diff. T-Stat
Pct. Female	0.41	0.38	1.00
Pct. Age 18-20	0.05	0.05	-0.07
Pct. Age 21-30	0.20	0.22	-0.55
Pct. Black	0.12	0.08	1.66
Pct. Other	0.03	0.02	0.68
Pct. Hispanic	0.15	0.14	0.66
Pct. Less than HS Educ.	0.16	0.12	1.53
Pct. Some College Educ.	0.28	0.30	-1.69
Pct. College Grad. Educ.	0.28	0.28	-0.24
Avg. HH Income	84.27	85.74	-0.35
Pct. Student	0.09	0.08	0.75
Avg. Liquor Drinks	36.97	24.79	3.17^{**}
Avg. Beer Drinks	22.00	18.18	2.24^{*}
Avg. Wine Drinks	13.68	8.74	3.55^{***}
Avg. Magazine Issues	258.16	176.98	2.81^{**}
Avg. Hours TV per Week	54.98	45.45	2.38^{*}
Mkt. Share in Type	0.06	0.22	-4.91^{***}
Market Leader	0.07	0.55	-4.82^{***}
Magazine Spending in 1996	1.97	3.88	-1.33
Observations	72		

Table 2.4: Number of Brands, Pct. Starters, Pct. Switchers, and First Movers by Spirits Type

Spirits Type	# Brands	% Started	% Switched	First Mover Brands
Blended Whiskey	1	0	0	Seagrams 7
Bourbon Whiskey	8	.25	.25	Jack Daniels
Canadian Whisky	4	.25	.25	Crown Royal
Cordials	18	.39	.39	Kahlua
Gin	5	.4	.4	Bombay
Irish Whiskey	3	.33	.33	Jameson
Rum	5	.6	.6	Captain Morgan
Scotch Whisky	11	.36	.36	Chivas Regal
Tequila	5	.4	.2	Jose Cuervo
Vodka	12	.5	.33	Grey Goose, Smirnoff

Table 2.5: Estimates of Relationship Between Initial Brand Characteristics and TV Advertising

	(1) Started TV	(2) Switched to TV	(3) Moved to TV First
Pct. Age 18-20	-1.5109	-2.1624	1.2573
	(2.0441)	(2.0809)	(1.6880)
Pct. Age 21-30	0.9514	0.6737	-0.3348
	(0.7691)	(0.6557)	(0.3908)
Avg. Liquor Drinks	0.0182	0.0323	0.0080
	(0.0226)	(0.0280)	(0.0235)
Avg. Beer Drinks	-0.0353**	-0.0508*	-0.0053
	(0.0151)	(0.0236)	(0.0159)
Avg. Wine Drinks	-0.0031	-0.0141	-0.0147
	(0.0371)	(0.0348)	(0.0416)
Avg. Magazine Issues	0.0000	-0.0010	-0.0005
	(0.0032)	(0.0029)	(0.0013)
Avg. Hours TV per Week	-0.0123	-0.0074	-0.0007
	(0.0167)	(0.0132)	(0.0059)
One Firm Concentration Ratio	0.3779	0.3611	-0.0180
	(0.4462)	(0.7040)	(0.4353)
Two to Four Firm Ratio	-0.5458	-0.4167	-1.5259
	(1.3063)	(1.5269)	(0.8483)
Mkt. Share in Type	1.8594	2.2860	0.2244
	(1.6125)	(1.6629)	(1.6598)
Sales of Brand (000 Cases)	-0.0429	-0.0036	0.1014**
	(0.0741)	(0.1162)	(0.0428)
Total Sales of Type (000 Cases)	0.0013	0.0013	-0.0009
	(0.0012)	(0.0013)	(0.0007)
Market Leader	-1.2455*	-1.4341*	-0.7381
	(0.6428)	(0.6482)	(0.8223)
Magazine Spending in 1996	0.0386	0.0143	-0.0306*
	(0.0328)	(0.0449)	(0.0162)
Market Leader x One Firm Ratio	-2.2350	-2.4240	-0.3751
	(1.7891)	(1.7122)	(1.9474)
Market Leader x 2-4 Firm Ratio	5.8900**	6.6821***	4.0710
	(1.8631)	(1.9092)	(2.3496)
Observations	72	72	72
R^2	0.565	0.505	0.507

Standard errors clustered at the spirits category.
Models also contain additional demographic controls (gender, race, education, and student status).
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.6: Simulated Relationship between Various One and Two to Four Firm Ratios and Predicted Probability of Ever Using Television Advertising for Market Leading Firm

		One Firm Ratio				
		10	20	30	40	50
2-4 Firm Ratio	20	.33	.3	.27	.25	.22
	30	.78	.75	.72	.69	.67
	40	1.22	1.2	1.17	1.14	1.11

Table 2.7: Simulated Relationship between Various One and Two to Four Firm Ratios and Predicted Probability of Ever Using Television Advertising for Firms Ranked Two to Four

		One Firm Ratio				
		10	20	30	40	50
2-4 Firm Ratio	20	.75	.85	.96	1.06	1.17
	30	.28	.38	.49	.59	.7
	40	-.19	-.09	.02	.12	.23

Table 2.8: Simulated Relationship between Various One and Two to Four Firm Ratios and Predicted Probability of Ever Using Television Advertising for Firms Ranked 5+

		One Firm Ratio				
		10	20	30	40	50
2-4 Firm Ratio	20	.44	.54	.65	.75	.86
	30	-.03	.08	.18	.29	.39
	40	-.5	-.39	-.29	-.18	-.08

Table 2.9: Estimates of Relationship Between Brand Characteristics and TV Advertising

	(2)	(1)
		TV Advertising Spending
Effect	Main	× Period
Pct. Age 18-20	-2.6128*	0.4420
	(1.5170)	(0.5332)
Pct. Age 21-30	-0.2881	0.2369
	(0.9565)	(0.2418)
Avg. Liquor Drinks	0.0205	-0.0047
	(0.0151)	(0.0056)
Avg. Beer Drinks	-0.0031	-0.0027
	(0.0151)	(0.0052)
Avg. Wine Drinks	-0.0431	0.0220*
	(0.0432)	(0.0125)
Avg. Magazine Issues	0.0012	-0.0003
	(0.0012)	(0.0004)
Avg. Hours TV per Week	-0.0082	0.0037**
	(0.0071)	(0.0016)
Mkt. Share in Type	-1.3368	0.3635
	(1.9110)	(0.7608)
Sales of Brand (000 Cases)	0.2867*	0.0162
	(0.1530)	(0.0462)
Total Sales of Type (000 Cases)	-0.0044**	0.0014*
	(0.0021)	(0.0008)
One Firm Concentration Ratio	1.4638	-0.3545
	(0.9018)	(0.3068)
Two to Four Firm Ratio	-4.3631	2.0237**
	(2.6414)	(0.9793)
Market Leader	-2.2025	0.2829
	(1.6136)	(0.6462)
Market Leader x One Firm Ratio	0.4571	-0.1323
	(2.4925)	(0.7834)
Market Leader x 2-4 Firm Ratio	5.0671	-0.2759
	(6.5922)	(2.5078)
Magazine Spending in 1996	0.0068	-0.0043
	(0.0302)	(0.0064)
t-1 Own Spending on TV	1.4895***	-0.1321***
	(0.2379)	(0.0454)
Current Year Rival Spending on TV	-0.0419	-0.0017
	(0.0961)	(0.0197)
t-1 Rival Spending on TV	0.1358*	-0.0229*
	(0.0742)	(0.0132)
At Least One Firm Started	0.3685	-0.0736
	(0.3724)	(0.1167)
At Least One Firm Switched	-0.6225	0.2167
	(0.5857)	(0.1703)
Period (Years Since 2001)	-0.6183	-0.6183
	(0.4306)	(0.4306)
Period Squared	-0.0121	-0.0121
	(0.0150)	(0.0150)
Observations	573	573
R^2	0.796	0.796

Standard errors clustered at the brand level.
Models also contain additional demographic controls (gender, race, education, income, and student status).
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.10: Hazard Models of TV Advertising Initiation and Switching

	(1)		(2)	
	Start TV Advertising		Switch to TV Advertising	
Effect	Main	× Period	Main	× Period
Pct. Age 18-20	0.0627	-0.0532	-0.1491	0.0119
	(0.2748)	(0.0772)	(0.2301)	(0.0750)
Pct. Age 21-30	0.0884	0.0294	-0.0014	0.0352
	(0.1807)	(0.0387)	(0.1835)	(0.0387)
Avg. Liquor Drinks	-0.0066*	0.0017*	0.0012	-0.0004
	(0.0038)	(0.0010)	(0.0028)	(0.0011)
Avg. Beer Drinks	0.0004	-0.0010	-0.0030	-0.0000
	(0.0041)	(0.0013)	(0.0030)	(0.0011)
Avg. Wine Drinks	0.0111	-0.0018	-0.0010	0.0015
	(0.0067)	(0.0018)	(0.0069)	(0.0019)
Avg. Magazine Issues	0.0002	0.0000	-0.0001	0.0000
	(0.0003)	(0.0001)	(0.0003)	(0.0001)
Avg. Hours TV per Week	0.0011	-0.0004	0.0010	-0.0001
	(0.0015)	(0.0004)	(0.0014)	(0.0004)
Mkt. Share in Type	-0.0561	-0.0426	-0.0096	-0.0005
	(0.4050)	(0.1560)	(0.2895)	(0.1154)
Sales of Brand (000 Cases)	0.0678*	0.0013	0.0451*	-0.0071
	(0.0359)	(0.0102)	(0.0241)	(0.0084)
Total Sales of Type (000 Cases)	-0.0001	0.0001	-0.0003	0.0001
	(0.0003)	(0.0001)	(0.0003)	(0.0001)
One Firm Concentration Ratio	-0.0898	0.0054	0.0909	-0.0710
	(0.1926)	(0.0616)	(0.1340)	(0.0570)
Two to Four Firm Ratio	0.0155	0.1123	-0.0920	0.1257
	(0.3830)	(0.1367)	(0.3969)	(0.1252)
Market Leader	-0.6529	0.1077	-0.2568	0.0989
	(0.4110)	(0.1274)	(0.1770)	(0.0915)
Market Leader x One Firm Ratio	-0.4848	0.2186	0.0897	-0.0402
	(0.5860)	(0.1653)	(0.3862)	(0.1332)
Market Leader x 2-4 Firm Ratio	2.7513	-0.7017	0.3263	-0.1782
	(2.0132)	(0.5256)	(0.5167)	(0.2781)
Magazine Spending in 1996	-0.0155	0.0119***	-0.0058	0.0024*
	(0.0095)	(0.0026)	(0.0052)	(0.0014)
Current Year Rival Spending on TV	-0.0123	0.0011	-0.0040	0.0003
	(0.0158)	(0.0026)	(0.0161)	(0.0027)
t-1 Rival Spending on TV	0.0199	-0.0029	0.0206	-0.0037
	(0.0173)	(0.0027)	(0.0170)	(0.0027)
At Least One Firm Started	-0.0096	-0.0236	0.0035	-0.0201
	(0.0607)	(0.0357)	(0.0505)	(0.0231)
At Least One Firm Switched	0.0303	0.0039	-0.0428	0.0077
	(0.0822)	(0.0200)	(0.0709)	(0.0193)
Period (Years Since 2001)	-0.0253	-0.0253	-0.0100	-0.0100
	(0.0610)	(0.0610)	(0.0518)	(0.0518)
Period Squared	-0.0062**	-0.0062**	-0.0000	-0.0000
	(0.0030)	(0.0030)	(0.0036)	(0.0036)
Observations	465	465	516	516
R^2	0.234	0.234	0.109	0.109

Standard errors clustered at the brand level.
Models also contain additional demographic controls (gender, race, education, income, and student status).
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix A: Auxiliary Model Specifications

Table 2.11: Estimates of Relationship Between Initial Brand Characteristics and Ever Advertising on TV

	(1)	(2)	(3)
Pct. Age 18-20	-0.6353	-0.6222	-1.5109
	(2.0617)	(2.0661)	(2.0441)
Pct. Age 21-30	1.2630	1.2799	0.9514
	(0.8525)	(0.8556)	(0.7691)
Avg. Liquor Drinks	0.0052	0.0058	0.0182
	(0.0246)	(0.0258)	(0.0226)
Avg. Beer Drinks	-0.0210	-0.0230	-0.0353**
	(0.0165)	(0.0187)	(0.0151)
Avg. Wine Drinks	0.0002	0.0009	-0.0031
	(0.0410)	(0.0410)	(0.0371)
Avg. Magazine Issues	0.0011	0.0011	0.0000
	(0.0031)	(0.0032)	(0.0032)
Avg. Hours TV per Week	-0.0141	-0.0144	-0.0123
	(0.0165)	(0.0169)	(0.0167)
One Firm Concentration Ratio	-0.3962	-0.2704	0.3779
	(0.6908)	(0.6732)	(0.4462)
Two to Four Firm Ratio	1.4434	1.2806	-0.5458
	(1.1424)	(1.0789)	(1.3063)
Mkt. Share in Type	1.0734	1.6588	1.8594
	(1.0678)	(1.7677)	(1.6125)
Sales of Brand (000 Cases)	-0.1058	-0.1240	-0.0429
	(0.1076)	(0.0843)	(0.0741)
Total Sales of Type (000 Cases)	0.0009	0.0010	0.0013
	(0.0013)	(0.0013)	(0.0012)
Market Leader	0.0930	0.2551	-1.2455*
	(0.3432)	(0.4446)	(0.6428)
Magazine Spending in 1996	0.0521	0.0532	0.0386
	(0.0399)	(0.0389)	(0.0328)
Market Leader x One Firm Ratio		-0.8632	-2.2350
		(2.0766)	(1.7891)
Market Leader x 2-4 Firm Ratio			5.8900**
			(1.8631)
Observations	72	72	72
R^2	0.523	0.526	0.565

Standard errors clustered at the spirits category.
Models also contain additional demographic controls (gender, race, education, and student status).
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.12: Estimates of Relationship Between Initial Brand Characteristics and Switching to TV Advertising

	(1)	(2)	(3)
Pct. Age 18-20	-1.1673	-1.1542	-2.1624
	(2.3547)	(2.3377)	(2.0809)
Pct. Age 21-30	1.0294	1.0464	0.6737
	(0.7357)	(0.7374)	(0.6557)
Avg. Liquor Drinks	0.0177	0.0183	0.0323
	(0.0323)	(0.0337)	(0.0280)
Avg. Beer Drinks	-0.0348	-0.0369	-0.0508*
	(0.0269)	(0.0291)	(0.0236)
Avg. Wine Drinks	-0.0103	-0.0096	-0.0141
	(0.0395)	(0.0388)	(0.0348)
Avg. Magazine Issues	0.0001	0.0001	-0.0010
	(0.0029)	(0.0030)	(0.0029)
Avg. Hours TV per Week	-0.0094	-0.0096	-0.0074
	(0.0135)	(0.0140)	(0.0132)
One Firm Concentration Ratio	-0.5009	-0.3744	0.3611
	(0.8211)	(0.8745)	(0.7040)
Two to Four Firm Ratio	1.8190	1.6553	-0.4167
	(1.2836)	(1.2007)	(1.5269)
Mkt. Share in Type	1.4699	2.0584	2.2860
	(1.1055)	(1.7945)	(1.6629)
Sales of Brand (000 Cases)	-0.0774	-0.0957	-0.0036
	(0.1356)	(0.1143)	(0.1162)
Total Sales of Type (000 Cases)	0.0008	0.0010	0.0013
	(0.0015)	(0.0016)	(0.0013)
Market Leader	0.1054	0.2683	-1.4341*
	(0.3881)	(0.5308)	(0.6482)
Magazine Spending in 1996	0.0298	0.0310	0.0143
	(0.0499)	(0.0491)	(0.0449)
Market Leader x One Firm Ratio		-0.8678	-2.4240
		(2.0808)	(1.7122)
Market Leader x 2-4 Firm Ratio			6.6821***
			(1.9092)
Observations	72	72	72
R^2	0.449	0.452	0.505

Standard errors clustered at the spirits category.
Models also contain additional demographic controls (gender, race, education, and student status).
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.13: Estimates of Relationship Between Initial Brand Characteristics and Moving First to TV Advertising

	(1)	(2)	(3)
Pct. Age 18-20	1.8802	1.8715	1.2573
	(1.2823)	(1.3136)	(1.6880)
Pct. Age 21-30	-0.0965	-0.1077	-0.3348
	(0.4320)	(0.4441)	(0.3908)
Avg. Liquor Drinks	-0.0002	-0.0006	0.0080
	(0.0205)	(0.0194)	(0.0235)
Avg. Beer Drinks	0.0019	0.0032	-0.0053
	(0.0106)	(0.0107)	(0.0159)
Avg. Wine Drinks	-0.0116	-0.0120	-0.0147
	(0.0406)	(0.0425)	(0.0416)
Avg. Magazine Issues	0.0002	0.0002	-0.0005
	(0.0012)	(0.0012)	(0.0013)
Avg. Hours TV per Week	-0.0023	-0.0021	-0.0007
	(0.0059)	(0.0057)	(0.0059)
One Firm Concentration Ratio	-0.3826	-0.4661	-0.0180
	(0.4653)	(0.3261)	(0.4353)
Two to Four Firm Ratio	-0.3716	-0.2636	-1.5259
	(0.5744)	(0.6582)	(0.8483)
Mkt. Share in Type	0.4743	0.0857	0.2244
	(0.8608)	(1.6106)	(1.6598)
Sales of Brand (000 Cases)	0.0333	0.0453	0.1014**
	(0.0772)	(0.0593)	(0.0428)
Total Sales of Type (000 Cases)	-0.0009	-0.0010	-0.0009
	(0.0008)	(0.0006)	(0.0007)
Market Leader	0.4067	0.2991	-0.7381
	(0.2994)	(0.4743)	(0.8223)
Magazine Spending in 1996	-0.0197	-0.0205	-0.0306*
	(0.0212)	(0.0195)	(0.0162)
Market Leader x One Firm Ratio		0.5730	-0.3751
		(2.0248)	(1.9474)
Market Leader x 2-4 Firm Ratio			4.0710
			(2.3496)
Observations	72	72	72
R^2	0.470	0.472	0.507

Standard errors clustered at the spirits category.
Models also contain additional demographic controls (gender, race, education, and student status).
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.14: Estimates of Relationship Between Initial Brand Characteristics and Ever Advertising on TV by Market Rank

	(1) Market Rank 1	(2) Market Rank 2-4	(3) Market Rank 5+
One Firm Concentration Ratio	0.0873	0.0333	0.5289
	(1.3403)	(1.6818)	(0.3548)
Two to Four Firm Ratio	6.0079***	-4.1604*	-3.2881***
	(1.8182)	(2.0442)	(0.8749)
Sales of Brand (000 Cases)	0.0473	0.1197	0.5374***
	(0.0840)	(0.1130)	(0.0898)
Total Sales of Type (000 Cases)	0.0029	-0.0009	-0.0002
	(0.0031)	(0.0027)	(0.0004)
Mkt. Share in Type		1.1947	-4.2487
		(2.0149)	(2.7546)
Observations	10	18	44
R^2	0.557	0.366	0.162

Standard errors clustered at the spirits category.
Models also contain additional demographic controls (gender, race, education, and student status).
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.15: Estimates of Relationship Between Initial Brand Characteristics and Switching to TV Advertising by Market Rank

	(1) Market Rank 1	(2) Market Rank 2-4	(3) Market Rank 5+
One Firm Concentration Ratio	0.0873	-0.0669	-0.1848
	(1.3403)	(1.9040)	(0.1911)
Two to Four Firm Ratio	6.0079***	-3.8599*	-2.7197**
	(1.8182)	(1.8084)	(1.1328)
Sales of Brand (000 Cases)	0.0473	-0.0418	0.5491***
	(0.0840)	(0.1816)	(0.0936)
Total Sales of Type (000 Cases)	0.0029	-0.0004	-0.0005
	(0.0031)	(0.0023)	(0.0006)
Mkt. Share in Type		1.8820	-3.8077
		(2.6944)	(3.8078)
Observations	10	18	44
R^2	0.557	0.200	0.191

Standard errors clustered at the spirits category.
Models also contain additional demographic controls (gender, race, education, and student status).
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.16: Estimates of Relationship Between Initial Brand Characteristics and Moving First to TV Advertising by Market Rank

	(1) Market Rank 1	(2) Market Rank 2-4	(3) Market Rank 5+
One Firm Concentration Ratio	1.3321	0.1546	0.1136
	(1.1592)	(0.9491)	(0.3036)
Two to Four Firm Ratio	3.5878	-1.6415	-1.4238
	(2.6745)	(1.5326)	(1.1272)
Sales of Brand (000 Cases)	-0.0469	-0.0075	0.0780
	(0.1487)	(0.1345)	(0.0835)
Total Sales of Type (000 Cases)	0.0025	0.0002	-0.0011
	(0.0032)	(0.0020)	(0.0011)
Mkt. Share in Type		0.5399	-0.2175
		(0.9777)	(1.5655)
Observations	10	18	44
R^2	0.312	0.076	0.075

Standard errors clustered at the spirits category.
Models also contain additional demographic controls (gender, race, education, and student status).
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

CHAPTER 3

LIVE AND LEARN OR LEARN AND LIVE: DOES EDUCATION LEAD TO LONGER LIVES?

This research is joint with Dean Lillard (Cornell University and DIW).

3.1 Introduction

Numerous studies establish a large and positive correlation between health and education, but few convincingly show whether the relationship is causal. Among the many factors correlated with health, social scientists and policy makers often focus on education because the production of human capital is publicly subsidized and because education is more strongly correlated with health than with either income or occupation (Grossman and Kaestner, 1997). While education certainly plays a role in determining both income and occupation, it is also independently correlated with health. Further, education is positively correlated with health measured in several ways: mortality rates, morbidity rates, self-evaluated health, and psychological health indicators (Grossman, 2006).

The literature has advanced three mechanisms to explain the correlation between education and health. The first mechanism is that a person who gets more education learns how to produce health more efficiently. The second is that if a person is healthy he pays less (in effort and time) to invest in schooling. The third explanation is that unobserved factors affect both schooling and health. However, these explanations are not mutually exclusive.

Researchers have long recognized that it is difficult to statistically identify which direction causality runs between education and health. This difficulty arises because both education and health are largely endogenous. People choose their ultimate educational attainment and they also choose whether or not to engage in activities that affect their health.

Because health is so strongly correlated with education, social scientists and policy makers are keenly interested in understanding whether the two are causally related, in which direction the causality runs, and how strongly one predicts the other. The keen policy interest arises in part because governments separately target both health and education with so many public policies. Policy makers could better design policies related to both education and health if they had more and better evidence on whether education causes people to be in better health, whether better health causes people to get more education, or whether another "third-variable" explains both.

This paper revisits the link between education and mortality. Similar to prior studies, we study the effect of plausibly exogenous variation in education driven by compulsory schooling laws on adult mortality. However, we use combine two data unique data sources particularly well-suited to studying the causal effect of education on mortality. We use a new compilation of compulsory schooling laws that are measured with less error. We combine these policy data with the Panel Study of Income Dynamics using a detailed state of residence history and a schooling law assignment algorithm that more accurately match respondents to the laws they actually faced. We compare our instruments to those used by previous studies and find substantive differences. The PSID Death File provides information on the date and cause of death for PSID respondents. We use these death data to estimate 5- and 10-year probit models of mortality and continuous-time models of survival past age forty. For each model type, we estimate both naive models of the statistical relationship between education and mortality and instrumental variables models of the casual effect of schooling on mortality.

We confirm a strong statistical relationship between education and mortality but find no evidence that the relationship is causal. Our first-stage estimates show a modest, but statistically significant effect of compulsory schooling laws on educational attainment. We show this effect is concentrated in the cohorts born from 1901 to 1945. We compare our instruments to those used by previous studies and find ours predict a larger effect of compulsory schooling on educational attainment. We also find that this effect is concentrated in grades 6 through 12. Our results show a strong relationship between education and 5- and 10-year mortality and survival past age 40. However, our instruments only weakly predict education. While we find no statistically significant causal effect of education on mortality or longevity in our instrumental variables models, the weakness of our instruments makes these results largely uninformative. We also find no effect of education on mortality by causes related or unrelated to health behaviors, though these results are also mostly uninformative. Finally, we show evidence of a particularly strong correlation between post-secondary education and mortality, though little evidence exists that this relationship is causal.

The rest of the paper is organized as follows: In Section 2 we review the literature that links education and health and the literature that empirically tries to identify whether there is a causal relationship in either direction. Section 3 describes our data, instruments, and highlights how they are better than existing instruments. In Section 4 we describe the methods we use and specify our models. We present results in Section 5. In Section 6 we discuss our results and conclude with observations about directions future research might take.

3.2 Background

Theoretical Framework

To build a causal role for education, Grossman's (1972) model of health capital assumes that when a person gets more education his health production function changes and that he may also learn more about how health inputs are related to health outputs. This first assumption says that education creates a new production function that yields more health for a given level of inputs than the old production function yielded. Stated differently, education teaches a person how to combine a given set of inputs differently so they produce more health. Grossman (1972) refers to this assumed effect of education as productive efficiency. For example, suppose two individuals both exercise on a treadmill for the same amount of time. A more educated individual might exercise more and less intensively during the workout in a way that produces better cardiovascular health. A study of Roman Catholic nuns found that although all of the sisters had essentially the same health inputs those with at least a bachelor's degree had half the mortality rate of nuns with less education (Snowdon, Ostwald and Kane, 1989). Cutler and Lleras-Muney (2008) find that the relationship between health and education increases with increasing levels of education and suggest that education alters thinking and decision-making patterns.

Second, because education conveys information, it is plausible to assume that, when a person gets more education, he might also learn new information about the relationship between inputs and health. Thus, a more educated person will better allocate his fixed resources across health and non-health related inputs. Grossman (1972) labels this assumed effect of education allocative efficiency.

If a less educated person does not understand or understands imperfectly how health inputs translate into health outputs then he is more likely to choose and combine inputs inefficiently. For example, suppose two individuals spend the same amount on health inputs. Under the hypothesis of allocative efficiency, a more educated person will choose from the full (or a fuller) set of health inputs. He chooses to use each input until the ratio of the value of the marginal product of the last unit and its cost is equal across all inputs. By contrast, a less educated person will simply ignore some inputs (whose marginal product yields greater value than they cost). As a result, he either chooses an incomplete mix of health inputs and/or combines them inefficiently. Kenkel (1991) establishes a positive relationship between schooling, health knowledge and health behavior. He shows that the more educated smoke less, drink to excess less, and exercise more.

The literature has also recognized that causality could run from health to education–people in better health might produce more education. For example, it is clear that students miss fewer days of school when they are healthier. As a result, they probably produce more human capital in a given amount of time. And, because they complete a given level of schooling in shorter time, are more likely to progress to each subsequent level of education earlier in their life-cycle (Edwards and Grossman, 1979).

Finally, the literature recognizes that there may not be a causal relationship between education and health (in either direction). Schooling will be positively correlated with health if an increase in an unobserved variable (independently) causes a person to produce more schooling *and* more health. Fuchs (1982) finds a correlation between time preference and smoking and health status, suggesting that intertemporal preference may be relevant third variable.

Empirical literature on causal effect of education on health

In the literature that tries to identify whether education causes health (or visa versa), most studies use either a standard instrumental variables (IV) approach or a more specific regression discontinuity method. The standard IV approach is more common in the literature. Technically regression discontinuity is a type of IV analysis, and the two methods have similar strengths and drawbacks.

Both methods pose statistical challenges. The IV method requires that one find variables (instruments) that predict variation in educational attainment but that are also uncorrelated with health but for their effect on education. This requirement, the "exclusion restriction," represents a statistical challenge because it is difficult to find variables that both predict education and are arguably orthogonal to health conditional on education. Some of the first IV studies relied on instruments that invoked questionable exclusion restrictions. Berger and Leigh (1989) predict education with state of birth, per capita income, and state education expenditures as instruments. If, for example, state per capita income is correlated with state per capita spending on health, which likely affects the health of residents, state per capita income is an invalid instrument for education when predicting the effect of education on health. More recent studies use instruments that are easier to defend such as the geographic availability of colleges (Currie and Moretti, 2003), state high school graduation requirements and state policies used to award certificates of General Educational Development (GED)

(Kenkel, Lillard and Mathios, 2006), and compulsory school entry laws (Adams, 2002; Lleras-Muney, 2005; Chevalier, 2004; Black, Devereaux and Salvanes, 2004; Arendt, 2005; Chou et al., 2010).[1] The IV method estimates a local average treatment effect (LATE) of education, which gives the effect of education on health for individual's affected by the instrument and at the level of education affected by the rule. For example, if the distance to college only affects community college attendance for women, then the IV approach will estimate the effect of community college attendance on the health of women. This LATE likely will not generalize to other populations or level of education.

More recently studies have employed regression discontinuity (RD) designs. The design of these studies exploits a threshold, usually created by policy, that puts observationally similar individuals into different educational categories because of small differences in a continuously distributed variable. For example, individuals graduate from high school if they score above some level on a high school graduation test and fail to graduate if they score less than the required amount. A set of studies uses school exit laws—laws that specify the age (or grade) an individual must complete before he may leave school—to compare health of the cohort that was first required to meet the (usually) higher age (grade) to the cohort just before the change was implemented (Chou et al., 2010; McCrary and Royer, 2011; Clark and Royer, 2010). The RD approach also estimates a LATE of education: the estimated effect is specific to the cohorts who were affected by the changed policy in a narrow window around the years just after the policy change and may not generalize to other cohorts. Also, the RD design exploits some rule that assigns different levels of education based on a continuously distributed variable (such as a test score threshold). The LATE gives the effect for individuals in the immediate vicinity of the threshold and may not generalize to those who are far from the cutoff. For these reasons, RD may provide also provide limited insight.

Our focus on mortality and our use of variation in compulsory school entry laws is most similar to Lleras-Muney (2005). She uses US census data to estimate whether education affects adult mortality and uses variation across states and time in compulsory school entry laws as instruments. Since the U.S. census data do not contain information about the death of respondents, the author imputes a mortality rate. To do so she aggregates the data at the count of the population in born in a given state and year and creates a pseudo-panel of that state-year birth cohort across decennial censuses. She then constructs the mortality rate as change in the population of a given birth cohort between censuses divided by the base year

[1] Angrist and Krueger (1991) first used compulsory schooling laws as instruments to study the returns to education.

population. To predict education, Lleras-Muney (2005) uses variation across states and over time in the compulsory school entry laws that applied to each cohort. She finds that increases in education result in statistically and economically meaningful reductions in the probability of mortality in ten-year age groups. The evidence suggests that education does indeed cause mortality to decline. However, because mortality is imputed, Lleras-Muney (2005) can only study 10-year mortality rates and cannot analyze the effect of education on mortality from different causes. Additionally, Mazumder (2008) finds the results of Lleras-Muney (2005) are sensitive to the inclusion of state-specific time trends.

Our study advances the literature in several ways. We use instruments (on compulsory school entry laws) that are richer and better measured than those used in recent studies because they are taken directly from state statutes. Our analysis also improves the first-stage prediction because we develop an algorithm that assigns each PSID respondent to a state of residence in each year of life. That algorithm reduces the measurement error in the assignment of the instrument we use to predict educational attainment. We do not rely on imputed mortality but use data on individual mortality that is drawn from the National Death Index. Those data are linked to respondents to the Panel Study of Income Dynamics and so we are able to model an individual's mortality risk over a large part of his life. The rich demographic data, measured over the life-course of PSID respondents allows us to control for individual heterogeneity that is typically subsumed in the error term in studies that use aggregate data. The PSID data also improve over existing literature because the PSID sample spans multiple birth cohorts that span a broad range of calendar years. Finally, we also take advantage of the inter-generational design of the PSID to link parents with children. This survey design allows us to control for the education of a person's parents, which may be an important variable omitted from previous studies.

3.3 Data

We use data from the Panel Study of Income Dynamics and the PSID Death File from 1968-2007. The PSID began in 1968 with a sample of 4,802 families containing 17,807 individual respondents. The survey was administered annually until 1997, when the survey became biannual. Respondents who moved away from one of the original families were followed to their new residence and their new coresidents were added to the survey, which expanded the survey to over 8,200 families in 2007. The 1968-2007 Death File contains information about all known deaths of PSID respondents. When possible, deceased PSID respondents were matched to the National Death Index archived by the National Center for Health Statistics to obtain up to five causes of each death. The 1968-2007 Death File contains date of death information for 5,712 decedents and the cause of death for 4,061 of these respondents. The cause of death is missing for 1,651 deaths that occurred prior to 1979, the year the National Death Index began keeping records. This implies that our analyses on the subset of deaths for which we observe a cause can only represent mortality after 1978.

We include controls for gender, race (black, Hispanic, and other), parent's highest education (less than high school, high school graduate, some college, and college graduate or higher)[2], and year of birth. We measure education by the highest grade level each respondent reports completing. Due to survey limitations, our measure of highest grade completed is top-coded at grade 17 (at least a year of postgraduate work). Since we analyze respondents that are 45 years or older, we are reasonably certain that we measure completed education.

We limit our total sample to individuals that have valid demographic, own education, and parents' education data and have valid survey responses at age 45 or older. This eliminates 900 PSID respondents known to have died before age 45. We further eliminate 935 decedents and 3,052 non-decedents that have no recorded parental education. Our final sample consists of 11,255 individuals, 3,030 of whom we observe dying. We further limit most of our analyses to 6,145 of these individuals born between 1901 and 1945.

Table 3.1 statistically describes characteristics of the sample from the survival models. The average person in the sample completed 11.1 years of education and lived in a state that required him to attend school for 8.7 years. Educational attainment among the parents of our sample matches well with comparable measures estimated from other nationally representative data for the same cohorts (Stoops, 2004): 78 percent have parents with less than a high school education and only 6 percent have parents with a college degree.

3.3.1 Dependent variables - measures of death

Our main dependent variables measure how long a person lives (in months) after his fortieth birthday and the risk of mortality in 5- and 10-year age spans, conditional on having survived to the immediately previous age. Using data on each person's month and year of birth, we count the months between a person's fortieth birthday and either their last interview date (if living) or the date of their death (if deceased). The longevity variable will be censored

[2]We define parent's highest education as the attainment of the highest educated parent or the education level of a single parent if data are only available for either the mother or father.

for two types of people. It is left-censored for respondents who were older than forty in the survey year that their data are first measured by the PSID.[3] The longevity variable is right-censored at the date of last survey for respondents who are still alive or who dropped out of the survey. While the non-responders can include people who have died without the knowledge of the PSID, the number of such people is likely to be small because of a significant effort by the PSID to track non-responders and link them to the National Death Index.

We use the longevity measure in two ways. We analyze it directly as a dependent variable in a continuous-time survival model. We also use it to construct variables that indicate whether a person died within 5- or 10-year age span. We define a sample as everyone who survived to reach a particular age and, conditional on having done so, set the value of this variable equal to one if the respondent died in the given age range. The variable equals zero if the respondent survived (and was in the PSID sample) at that age just before the lower bound of the age range and the person did not die during the ages in the mortality age range. For example, we select all people who had a valid interview and were alive at age 49. We then identify all people who survived (and had valid data) at age 55. We measure the 5-year mortality risk as the sum of people in this group who died at age 50 through 54 relative to the total number of PSID respondents at risk to die in this age span.

Our data on the deaths of PSID respondents come from a special file of the PSID.[4] The data include, for all PSID respondents who have died, the month and year a person died and, when possible, the cause of death. To collect these data, the PSID linked each decedent to the National Death Index. The PSID Death File contains detailed information on the cause of death for 71% of decedents. For this group, the data include codes that detail the primary cause of death. These codes, taken from the death certificate correspond to the codes used by the International Statistical Classification of Diseases and Related Health Problems (ICD). Those codes have been revised 10 times since they were first established. Depending on the year the person died, the cause of death codes in the PSID Death File correspond to either Revision 9 codes (ICD9) or Revision 10 codes (ICD10).

[3] This group is unlikely to be large. The PSID defines respondents as "original sample unit members" (OSUMs) and "other family unit members" (OFUMs). Original sample members consist of living in a household surveyed in the base year (1968) and all of their offspring. Everyone else is an OFUM. It consists of OSUMs who were over 45 years old in the base year (1968) interview, OFUMs who joined the household of an OSUM in a later year, (these are usually parents or other relatives who moved into the home of their adult child), and OFUMs who enter the PSID sample because they marry a PSID sample member (after age 45).

[4] Due to the sensitive nature of the data, researchers must apply to access the PSID Death File data. See http://simba.isr.umich.edu/restricted/Mortality.aspx.

We use these data to separate deaths into one of two types - deaths from causes that could plausibly be affected by consciously chosen health behaviors and deaths from all other causes. The diseases we classify as related to behavior are cancers of the mouth, throat, esophagus, and lung; emphysema; diabetes; and heart disease. We use this simple classification to test whether education has a larger (or smaller) correlation with deaths related to health behaviors and whether education alters the probability of death from those causes. We recognize that many deaths we label as due to "behavioral" causes may not actually be caused by behaviors. Similarly, many deaths we label as "non-behavioral" may be strongly linked to behavior (e.g., suicides and drunk driving accidents). We make this distinction as a rough categorization of causes of death.

3.3.2 Compulsory schooling laws

We improve on the existing literature in part because we use data on compulsory schooling laws that are measured with less error. We use data compiled by Jennifer Gerner and Dean Lillard at Cornell University and labeled as the Compulsory Schooling Law (CoSLAW) database. The data cover the complete history of compulsory schooling laws for each state. The data were transcribed directly from printed laws from every legislative session in every state since 1852.[5]. While other compulsory schooling law databases exist, the data they contain are typically drawn from cross-sectional snapshots of state laws taken every 5 years or so.[6] The CoSLAW data include the exact date specific laws changed, which specific sections changed, and the date the laws took effect. As a result, the CoSLAW data are more accurate relative to other compilations. The database includes the earliest age at which parents may enroll their children in public school, the age by which they must enroll their children in school, and the youngest age a student may drop out of school.

In addition to using more accurate compulsory schooling law data, we also reduce measurement error because we better match a person to the law that he actually faced. In several previous studies, the data often identify the state in which each respondent was born but do not observe when they moved (if they live in a different state). For example, Angrist and Krueger (1991), Acemoglu and Angrist (2000) and Lleras-Muney (2005) assign respondents

[5]The Cornell University Law School Library holds one of the most complete collections of state statutes and state session laws in the United States

[6]One notable exception is Goldin and Katz (2011). They identify states that change their laws between these cross-sectional snapshots and reference legislation for the date of the change. They compare these data to those used in Acemoglu and Angrist (2000) and Lleras-Muney (2005).

the compulsory schooling laws in effect in the respondent's state of birth but assign the laws that were in effect in the year the respondent turned 14. This assignment rule suffers from two types of measurement error. First, it assigns the incorrect laws to all respondents who moved across state lines between their birth and the year they were required to enter school.[7] Second, this rule assigns the incorrect laws when a state changed its laws between the year a person was required to enter school and the year he turned 14. Mazumder (2008) attempts to improve on Lleras-Muney (2005) by matching each individual to the entry age mandated in his state in the year he turned 8. While an improvement, this approach may also assign the incorrect ages for those in school when their state changes its laws.

To reduce the measurement error from these two sources, we develop an algorithm to more accurately identify the state where a person resided in every year of life. This algorithm produces a state of residence history that allows us to more accurately match a person to the state in which she resided when the laws applied. Our place of residence matching algorithm uses all information available in the PSID: reported birth state, reported state the respondent "grew up" in, the birth states of younger siblings, and the state resided in for each valid survey wave. We use information on the month and year a respondent moved to their current address from each survey wave to assign the date of any cross-state move. We further impute the date of any other cross-state moves by using the age-specific probabilities of cross-state moves to assign the move year as the year where the cumulative probability of having moved during or before that year is fifty percent. We describe our state residence assignment algorithm further in Appendix A. Though we use more information on an individual's state of residence and moving history than previous studies, we do not observe a complete residence history and must impute a respondent's state for years prior to them entering the PSID sample.

To match laws to PSID respondents, we use the more accurately measured compulsory schooling laws data and each respondent's state of residence history. Using the residence history of each person, we merge the laws that he faced in every year (in the state where he resided in that year). We then identify the year the law applied. That is, for each person, we check whether a person is younger or older than the age at which his state's government compels him to enter school. In our algorithm, we exploit the fact that maximum permitted entry age laws typically specify a date (often the first day of classes) by which a child must be

[7] Endogenous mobility, moving to a state because of its compulsory laws or other factors correlated with them, would endanger our identification strategy. Most of our models include fixed effects for the state an individual lived in when he was fourteen. However, while rare, endogenous moving at young ages could be correlated with both a state's education laws and later life health, invalidating our exclusion restriction.

a certain age to be compelled to attend school. The CoSLAW data include this information. When we match, we use the respondent's date of birth to determine the age he is on the date specified in law. In the year a person is going to turn the maximum permitted entry age dictated by the state, we assume that state's law was a binding constraint. We assign that state's maximum permitted entry age as the age that respondent faced. We follow the same rule to assign each person the minimum permitted dropout age that was in effect in the year he first became eligible to (legally) leave school. The only difference is that minimum permitted dropout age laws typically only specify an age at which a student may leave school.[8] We use these two ages to create a variable that measures the number of years of compulsory required by each state (the mandated exit age minus the mandated entry age).

As detailed above, we take three steps to reduce measurement error due to the misassignment of compulsory schooling laws to individuals: the improved CoSLAW database, a better law assignment algorithm, and a comprehensive state of residence history. While our approach is theoretical improvement, we demonstrate here that the database and assignment algorithm create real differences in the assignment of compulsory schooling laws to individuals. However, given the lack of early-life state of residence information in the PSID for the cohort we study, we observe very few respondents moving prior to age 18. Consequently our state of residence history negligibly differs from the assumption an individual lives in their birth state for the entirety of his primary and secondary education. To test our database and law assignment algorithm, we create a subsample of individuals born 1904 to 1943 to whom we can assign both our CoSLAW instruments and those used by Acemoglu and Angrist (2000). We compare the differences in these law databases, holding the assignment algorithm constant. We then compare the differences in assignment algorithms, holding the law database constant.

Figure 3.2 shows the distribution of differences in assigned years of compulsory schooling from the CoSLAW database compared to the data used by Acemoglu and Angrist (2000) (i.e., the assigned years of schooling from our database minus the years from Acemoglu and Angrist (2000)). For both data sources, we assign the number of years required in a person's state of residence in the year he turned fourteen. The two data sources differ in their assignment for 995 of 5,615 individuals (18 percent). Our CoSLAW data assign an average of 0.63 additional years of education. A large majority of the differences are an assignment of one year more

[8]Our assignment rule assumes that parents do not move across state lines to subject their child to different entry or exit ages.

or less of compulsory education. These differences are largely due to disagreements in the timing of law changes. Because we observe the exact year a state adopted a law rather than a snapshot of the laws in each state every few years, we observe changes in laws earlier and are more likely to assign slightly more years of compulsory education. While these differences may seem trivial, most within-state variation in compulsory schooling, after a state adopts compulsory education, is small. As mentioned above, states seldom change laws, and those changes are typically increases of only a year or two. Properly assigning the date of the changes may be crucial to maximizing the use of this small amount of variation.

In Figure 3.3 we compare the number of years of compulsory education assigned by our algorithm to those assigned by simply using the laws in effect in the year a respondent turned fourteen. We use the CoSLAW database for both assignments, holding the source of the laws constant. Again, we find few differences in assignment: 515 of 5,615 (9 percent) differ. On average, our algorithm assigns 0.81 fewer years of compulsory education. The majority of differences are due to our algorithm assigning one fewer year of compulsory schooling. This typically occurs when a state lowers their compulsory entry age after a person was forced into school at the previous, older age but before the person turned age fourteen. We believe our algorithm is a better assignment of the law that person actually faced.

These analyses provide some evidence that our assignment algorithm and improved database more accurately assign the compulsory schooling ages required by a literal reading of state laws. However, to be effective these laws must be widely known and enforced. Given the often weak enforcement of compulsory schooling laws, it seems possible that enforcement lags behind the passage of stricter laws (Lleras-Muney, 2002). In this case, the cohorts we "more accurately" assign to newly passed laws may actually largely face the old laws due to their state's lack of resolve or preparedness to enforce the newer rules. Therefore, our theoretically more accurate assignment may not necessarily be any better, and could be worse, than previous measures.

3.4 Methods and Identification

General model

Grossman's model of health capital can be written in the general form:

$$HS_{it} = f(I_{it}, E_i, V_{it}, HS_{it-1}), \tag{3.1}$$

where HS_{it} is present period health stock, E_i is education, and V_{it} is a vector of prices of health-related goods (Grossman, 1972). Grossman assumes that an individual dies when his health stock falls below a specified value. Of course we do not observe a person's health stock directly. However, because a person's health stock is negatively related to the probability of death we can use the standard latent variables approach. We assume health stock is a function of individual characteristics, education, parent's education, and unobserved state factors.

Naive models of the relationship between education and mortality

We first estimate the correlation between education and mortality without accounting for the potential endogeneity of education. This potential endogeneity implies that there are unobserved factors in the error term that are correlated with both education and mortality. Because we do not account for this endogeneity in these models, this spurious correlation will be expressed by biasing our estimates of the effect of education on mortality (e.g., β_1 or γ_1). The first set of models estimate the probability an individual dies in a five- or ten-year age span, conditional on reaching the start of that span. For example, we model the probability of dying at age 45 to 49 (inclusive), conditional on having survived to age 44.[9] Generally, we estimate the probit model:

$$Pr(D_{ia} = 1) = \Phi(\beta_0 + \beta_1 E_i + X_i\beta_2 + P_i\beta_3 + W_s\beta_4), \tag{3.2}$$

where D_{ia} indicates the death of individual i in age span a, E_i is individuals i's education measured in the highest grade she has completed, X_i is a vector of (time invariant) individual characteristics, P_i a set of indicators the highest education of individual i's parents, W_s is a vector of 50 indicators that identify the person's state of residence at age 14.

We also estimate parametric, continuous-time survival models of the length of time past age 45 an individual lives. We estimate an accelerated failure time model using a Weibull distribution. This specification allows for the intuitive form of a semilog model of survival

[9] The sample is restricted to respondents who were younger than the lower bound of the age span when they entered the PSID sample. We include everyone who subsequently died (during the age span) and for whom we can determine their life status at an age older than the upper bound of the age span. Our rule excludes individuals who did not respond to the survey during and after the age span because we cannot verify their life status.

time:

$$ln(T_i) = \gamma_0 + \gamma_1 E_i + X_i\gamma_2 + P_i\gamma_3 + W_s\gamma_4 + \sigma u_i \tag{3.3}$$

where T_i is the amount of time (in months) that an individual i lives after he turned forty, σ is a shape parameter related to how the hazard rate increases (or decreases) over time, and u_i is an error term with a Type-1 extreme value distribution (Jenkins, 2005).

Instrumental variables approach

Similar to prior IV studies that use compulsory schooling laws as instruments, we estimate a first-stage equation using ordinary least squares. To predict the number of years of education an individual attains, we assume that attained education varies linearly with the years of schooling their state required them to attend:

$$E_i = \delta_0 + C_{st}\delta_1 + X_i\delta_2 + P_i\delta_3 + W_s\delta_4 + \eta_i, \tag{3.4}$$

where E_i is individual i's education measured as the highest grade she has completed, C_{st} is the years of schooling that state s required individuals to complete in year t, and η_{it} is the error term. X_i, P_i and W_s are defined as above.

In addition to this model, we estimate a set of auxiliary regressions to explore the impact of compulsory schooling laws. We estimate models by cohort groups to confirm the time period in which these laws had the greatest impact. We also evaluate the specific improvements we make to our instrument and, using the PSID, compare our instrument to those used in previous studies. We also study whether compulsory schooling laws differentially affect the probability of attaining different levels of education. In these models, the dependent variable for a particular education level equals zero if the person dropped out of school before completing that many years of education and it equals one if she completed at least that many years. For example, the indicator for the model of completing 8 years of education equals 0 for people who dropped out of school in grade 7 or earlier and equals 1 for people who stopped attending school after completing grade 8 or continued to complete higher grades (e.g. the indicator would equal one for grade 9 dropouts as well as high school graduates).

Rather than the more common approach of two-stage predictor substitution (2SPS), we use two-stage residual inclusion (2SRI), also referred to as two-stage conditional maximum likelihood (2SCML), to instrument for education in our second-stage (Rivers and Vuong, 1988). Though these two methods are mathematically equivalent in a linear model (i.e., two-stage least squares), Terza, Basu and Rathouz (2008) show that for a broad set of nonlinear models 2SRI is consistent, but 2SPS is not. Both approaches require an identical first-stage. However, 2SRI does not replace the endogenous regressor in the second-stage with a predicted value. The 2SRI approach retains the endogenous regressor and adds the first-stage residual as an additional regressor.

Intuitively, 2SRI works by separating out the "endogenous" component of an individual's education. As mentioned above, the naive approach will produce biased estimates of the effect of education because education may be correlated with unobserved factors that also correlate with health. While we cannot directly observe the correlation of education and these unobservables, we can consistently estimate this effect with a valid instrument. As a simple model, assume that education is the function of compulsory schooling laws and a vector of unobserved variables correlated with both education and health (Z_i):

$$E_i = C_{st}\alpha + \psi Z_i. \tag{3.5}$$

If we can consistently estimate α, we can consistently estimate ψZ_i, which is the component of education orthogonal to the component of education driven by compulsory schooling laws:

$$\psi Z_i = E_i - C_{st}\alpha. \tag{3.6}$$

Including ψZ_i in the second stage eliminates the omitted variables problem by controlling for the "endogenous" component of education. The 2SRI approach also allows for a simple test of the exogeneity of education.

Given the first-stage outlined above and the 2SRI method, we then estimate second-stage models of five- and ten-year mortality:

$$Pr(D_{ia} = 1) = \Phi(\beta_0 + \beta_1 E_i + X_i\beta_2 + P_i\beta_3 + W_s\beta_4 + \beta_5\hat{\eta}_i) \tag{3.7}$$

$$\hat{\eta}_i = E_i - \hat{E}_i, \tag{3.8}$$

where $\hat{\eta}_i$ is the predicted residual from the first-stage. We estimate the first-stage for each individual model so that we use the same sample in both stages. We also estimate equivalent instrumental variables survival models:

$$ln(T_i) = \gamma_0 + \gamma_1 E_i + X_i\gamma_2 + P_i\gamma_3 + W_s\gamma_4 + \gamma_5\hat{\eta}_i + \sigma u_i. \tag{3.9}$$

While the 2SRI estimator is mathematically consistent, the standard requirements of IV estimation apply. The instruments must be sufficiently strong predictors of education, and they cannot be correlated with the error term in the second-stage.

We also estimate more common 2SLS models (i.e., linear probability models) of 5- and 10-year mortality. These models have drawbacks in models of limited-dependent variables, particularly their assumption of a linear relationship between the probability of the dependent outcome and the independent variables and predicted probabilities less than zero and greater than one. However, they are simple to compute and their coefficient estimates are much easier to interpret.

Further tests of the relationship between education and mortality

We estimate several other models to explore more specific channels through which education might affect mortality. In the above models we have assumed that the probability of death varies in the same way for a unit change in education regardless of whether education changes by one unit from a base of 9 years or a base of 12 years. However, it is reasonable to suppose that a person might learn more information specifically relevant to health behavior in some years in school than in others. It is also reasonable to suppose that a non-causal relationship between education and mortality may be stronger for different levels of schooling. These suppositions suggests that the relationship between education and risk of mortality will vary with completed years of schooling. To explore this hypothesis, we split educational attainment into seven discrete categories: 1 to 6 years, 9 to 11 years, 12 years (a high school degree), 13 to 15 years (some college), 16 years (a bachelor's degree), or more than 16 years (some graduate school). We then estimate a survival model using indicators of these categories as covariates, rather than a continuous measure of education. Our auxiliary analyses and those of others (Acemoglu and Angrist, 2000; Lleras-Muney, 2002) show that compulsory schooling laws predominately affected secondary education. Because we lack instruments for each level of education, we can only estimate how the association between education and mortality varies with the level of completed education.

Finally, we exploit the data on cause of death to investigate whether educational attainment predicts the risk of death from specific causes that are more directly related to health behaviors. We estimate probit and 2SRI 10-year mortality models of death from causes related to health behaviors.[10] Here we define the dependent variable to equal one if a person died during a 10-year age span from a specific cause more directly related to a health behavior and zero if he did not die.[11] We also estimate comparable models of death from non-health behavior related causes.

3.5 Summary of Instruments and First-Stage Models

The contribution of this study largely rests on the validity of our instruments for education. Therefore, we first explore the variation in compulsory schooling laws and study how our instruments, created using the CoSLAW database and assignment algorithm, differ from compulsory schooling law instruments used in other studies.

The relevance of compulsory schooling laws as instruments depends on the laws strongly affecting educational attainment. Of particular importance is which birth cohorts compulsory laws affected. Lleras-Muney (2002) argues that these laws most strongly affected those born between 1900 and 1925. She cites evidence that poor enforcement for cohorts born prior to 1900 weakened the effect of the laws. Edwards (1978) shows little effect of the laws from 1940 to 1960, though this may be due to lack of variation in these laws during this time period. Table 3.2 reports estimates of the effect of compulsory laws for three different birth cohorts to confirm these results in the PSID. We find a strong, positive relationship between compulsory laws and schooling for the cohorts born from 1901 to 1945. However, we find a *negative*, statistically insignificant relationship for cohorts born prior to 1901 or after 1945. These results further confirm the conclusion that compulsory schooling laws primarily affected those born in the latter half of the 20th century. Therefore, we restrict all of our subsequent analyses to the 1901 to 1945 birth cohorts.

The validity of compulsory schooling laws as instruments further rests on their excludability in the second stage. As detailed above, most of our models include state fixed effects to account for time invariant state factors that may be correlated with both mortality and compulsory schooling laws. Most of our models also include parental education since parent's education may be correlated with both state compulsory schooling laws and their child's mortality. After including state fixed effects, the first-stage relies on within-state variation

[10] We detail our classification of health behavior related causes of death in our Data section.

[11] We exclude those who die from a non-health behavior related cause during this age span.

in compulsory laws to predict variation in educational attainment. If states seldom change compulsory laws, there is little variation to separate the time-invariant state effect from the effect of its laws. Figure 3.1 shows the distribution of the number of changes states made to their compulsory schooling laws between 1900 and 1960.[12] There few changes in compulsory laws over this time period: six states never change, seventeen states change only once, and twenty-eight states change two or three times. This implies that after including state fixed effects there is little remaining variation to estimate the effect of laws on schooling.

We report the results of our first-stage in Table 3.3. The first two columns present results from models without state fixed-effects. The second two columns report results from models with state fixed effects. Here the coefficients compare the average educational attainment of people in the same state before and after the policy changes. Columns (1) and (3) present results from models that do not control for parent's education. Columns (2) and (4) add three indicators of parental education (the reference category includes respondents whose highest educated parent completed only a high school diploma).

The results show that compulsory schooling policies affect years of completed schooling in intuitively plausible ways. There is a statistically significant increase in completed years when states require youth to stay in school another year. Results from the models with state fixed effects imply that compared to other cohorts attending school in the same state, the average youth born into a cohort that was required to attend one more year of schooling completes 0.22 more years of education. We find smaller, less precise point estimates of the effect of compulsory laws after adding state fixed effects. This result is consistent with a positive omitted variables bias from excluding time invariant state factors from the model. We also find that the estimated effect of compulsory laws does not vary much when one controls for parental education. Unsurprisingly, people whose parents are more educated complete more years of schooling. Relative to people whose parents finished high school, the average person whose parents completed a college degree get 1.7 more years of education. At the other end of the education distribution, people whose parents had dropped out of high school completed about 1.8 fewer years of schooling than people whose highest educated parent had completed high school.

In Table 3.4 we compare the first-stage results using six different instruments. In column (1), we use the data from Acemoglu and Angrist (2000) assigned at age fourteen and include their alternative minimum required years of schooling measure.[13] Column (2) omits the

[12] The laws in place over this time period affect the 1900 to 1945 birth cohort.

[13] Similar to Acemoglu and Angrist (2000), we assign the minimum of the minimum exit age minus the

alternate minimum years of schooling, making the measure more similar to those from our CoSLAW data. Column (3) uses the same instrument as column (2) but assigns laws based on our algorithm. In column (4), we use an alternative measure of compulsory schooling from Acemoglu and Angrist (2000) that equals the minimum age to obtain a work permit minus the maximum school entry age.[14] In column (5), we use our CoSLAW data assigned at age fourteen. Finally, column (6) includes our CoSLAW data assigned using our algorithm, our preferred specification used in all other models. Using the data from Acemoglu and Angrist (2000), we find that our assignment algorithm leads to smaller point estimates of the effect of compulsory laws on education (columns (2) and (3)). However, using the CoSLAW data, we find the opposite (columns (5) and (6)). The point estimates of the effect of the laws are larger. While the cause of this paradoxical result is unclear, the errors due to the less accurate law database and less accurate age fourteen assignment rule may cancel each other out. Comparing the law databases while holding the assignment rule constant (columns (2) and (5) and columns (3) and (6)), we find a larger effect of the laws using the CoSLAW data. However, the differences in the point estimates are not statistically significant. Though there are number of reasons these differences might occur, one potential explanation is that the point estimates using our data suffer from less attenuation bias due to law misassignment.

Table 3.5 replicates Table 3.4 without state fixed effects. There are less dramatic differences in the point estimates of the effect of schooling laws in these models. These models rely less on within-state variation in laws over time and more on cross-state variation. There is more agreement in the cross-state variation between these instruments than within-state variation, which largely relies on small discrepancies in the timing of law changes. This likely explains the closer point estimates.

As a final exploration of the effect of compulsory schooling, we study the relationship between compulsory schooling laws and the probability of completing a specific number of years of education in Table 3.6. The dependent variable in each of these linear probability models is a variable that takes a value of either zero or one. It indicates whether a person attained or exceeded x years of schooling. Over the period we study, the average person in the sample was required to be in school for 8.7 years. The results in Table 3.6 show that an additional year of compulsory schooling increases the probability of completing each level of schooling up to 11 years of education. Past 11 years, an additional year of mandated school attendance has no significant effect on the probability that a person attends post-

maximum entry age and the number of years of required schooling (often separately specified in state law).

[14]This measure is also used in Lleras-Muney (2002) and Lleras-Muney (2005).

secondary schooling. Further, the effect of compulsory schooling has the smallest effect on the probability of completing grade 6 (one more required year raises the probability of completing at least six years of schooling by 0.017 percentage points). An extra year of mandated schooling has a bigger effect on the probability of completing subsequent grade levels. Figures 3.4 plots the coefficients from the eleven regressions and demonstrates the sharp drop-off in the effect of compulsory schooling laws after grade 12.

3.6 Mortality Results

Five- and ten-year mortality models

We next turn to the analysis of mortality rates over 5- and 10-year intervals. Table 3.7 reports the number of people in our sample who were at risk to die in each of four ten-year age spans and the average mortality rate over the ten years. To explore whether compulsory school affects the probability of dying in these age intervals we first estimate naive models that treat education as if it were exogenously assigned. We report those results in Table 3.8. We then instrument for attained education using both 2SRI probit and 2SLS linear probability models. Table 3.9 presents results from 2SRI probit models. Table 3.10 presents results from 2SLS linear probability models.

Results in Table 3.8 replicate the well-established strong statistical association between education and mortality. Naive associations suggest that better educated people are less likely to die than less well educated people in every ten year age span. The association is strongest for people age 80-89 at -0.0191. In the next higher age category, those age 70-79, the association falls to -0.0102. This pattern is intuitively sensible if the mortality advantage for more educated people enjoy operates by reducing the risk of mortality from causes they can affect. For example, it is well-established that mortality rates of smokers and non-smokers begin to diverge around age 70 (Christopoulou et al., 2011). Under this sort of scenario one would expect the association between education and risk of death to decline after some age for two reasons. First, less educated people will comprise a smaller proportion of survivors. Second, the risk of death from exogenous causes rises with age. That is, at older ages people are more likely to die from causes that are unaffected or less affected by behavioral choices.

Table 3.9 and Table 3.10 present the results of our instrumental variables approach. However, these results provide little insight into whether this negative correlational relationship between education and mortality is causal. The F-statistics on the instruments show that they are very weak for all models and *extremely* weak for ages 60 to 69 and 70 to 79. Assuming our instruments are excludable from the second stage, a weak instrument will bias our estimates towards the naive results in Table 3.8. The results show that the effect of the instrumented education on mortality is not statistically different from zero. However, compared to Table 3.8, our second-stage point estimates are very imprecise with very large standard errors, and we cannot rule out improbably large or small effects. Due to the weakness of our instruments, the models do not provide any evidence to support or refute the hypothesis that education causes people to live longer.

To delve more deeply, we re-estimate the models substituting mortality risk in five- rather than ten-year age spans and reach the same conclusion. Table 3.11, 3.12, 3.13, and 3.14 report analogous results for models of 5-year mortality. These results are largely similar to the 10-year mortality results. Results from a naive regression presented in Table 3.12 shows a growing correlation between education and mortality as people age. Table 3.13 and 3.14 show a mix of insignificant positive and negative effects of education on mortality that are very imprecisely estimated. The instruments are generally weak in columns 5 to 8, and the standard errors of the point estimates increase by an order of magnitude compared to Table 3.12. In the four samples where the instruments predict educational attainment well, the estimated effects of education on 5-year mortality risk is generally negative but statistically insignificant. However, we do find a counterintuitive and significant *positive* effect of education on the probability of dying between age 45 to 49. Similar to the 10-year models, the analysis of the causal effect of education on 5-year mortality risk is inconclusive due to the weakness of our instruments.

Survival models

In Table 3.15 we report results from survival models that directly estimate whether education causes people to live longer. Recall that the dependent variable measures the number of months a person lives past his fortieth birthday. We estimate three sets of models that successively add controls. Results presented in columns (1), (3), and (5) are from models that treat education as exogenous. Results in columns (2), (4), and (6) instrument for exogenous variation in education associated with compulsory schooling. The first two models do not control for parental education. The last two models add state fixed effects. Given the semi-log form of the accelerated failure time model, the coefficients can be interpreted by the formula $\%\Delta T = 100[e^{\beta_i} - 1]$, where T is the survival time.[15]

Results in Table 3.15 are inconclusive, similar to the above results. If one ignores systematic differences in who gets education, one would conclude that education lengthens life. This association is robust to adding controls for parental education and state fixed-effects. The coefficients on years of education in columns (1), (3), and (5) range from 0.0137 to 0.0135 and always differ from zero at conventional thresholds of statistical significance. The coefficient in column (1) implies that an additional year of education is associated with living 1.2 percent longer past age 40. This correlation is modest but not insignificant in size. A white male born in the 1930's with 9 years of education can expect to live 5 months longer (75.4 years) than a similar man with only 8 years.

The association becomes positive, implying education increases mortality risk, when one accounts for systematic differences in educational attainment explained by compulsory schooling laws. However, this result is only statistically significant for the model with state fixed effects, where the instruments are the weakest. It is difficult to explain this result, and it may be due to the lack of remaining variation in the instruments. Additionally, these instrumental variables models are also estimated with much less precision than the naive hazard models. The evidence from the survival models provides no evidence that getting more education in years below high school affects length of life.

As with all studies using instrumental variables, our IV models predict the local average treatment effect (LATE) for individuals whom compulsory schooling laws force to get more education. Therefore, the LATE of our models is specific to those persons affected by the rules and the level(s) of education they then receive. Our first-stage analyses show that the effect of these laws is largest in grades 8-12. Combined with our IV results, this implies that we find no measurable effect of compulsory secondary education on mortality.

Results from supplementary analyses

We next report results from three supplementary analyses that tries to delve deeper into possible mechanisms by which education might causally affect mortality. We first divide deaths by the cause that was listed on the death certificate. In Table 3.16 we report results from models that analyze the 10-year probability of death from causes whose probability is known to vary with health behavior (versus people who did not die). In Table 3.17 we report

[15]For coefficients with an absolute value less than 0.25, simply multiplying the coefficient by 100 is a good approximation of the implied percentage change in survival time.

results from models that analyze the 10-year probability of death from all other causes. In both samples, we estimate naive and 2SRI models.

The results again point to the conclusion that exogenous variation in educational attainment does not affect mortality between age 50 and 69, regardless of whether the physicians attributed the death to causes that plausibly vary with health behaviors or not. As with our general 10-year mortality models, the standard errors of the 2SRI models are an order of magnitude larger than the naive probit models, implying that our coefficients are imprecisely estimated. Our instruments are again weak for the models of mortality above age 70, precluding any interpretation of the IV results.

All of the above results suggest that increases in educational attainment do not alter how long people live. However, all of these models are identified from differences in educational attainment that are predicted by variation in the compulsory schooling their state required them to complete. As we showed above, the compulsory schooling laws predict variation in education up through grade 12 but do not predict the probability a person completes education past high school. In light of these findings, it is possible that our measures of (exogenous) education fail to capture relevant variation in the true underlying human capital that affects health behavior. Given the typical high school curriculum, this conjecture seems plausible.

To shed some light on the possibility that mortality advantages start to accrue with post-secondary education, we estimate survival models where we include separate indicators for six categories of educational attainment. The survival of people with these levels of education is relative to those who graduated from high school. The results in Table 3.18 simply correlations and do not allow a causal interpretation. The model in column (2) adds control for parental education. The model in column (3) adds state fixed-effects.

Although the results only measure an association, they are consistent with the hypothesis that the human capital conveyed in college may affect mortality. They are also consistent with the hypothesis that those who graduate from college are healthier for reasons other than education. Results in column (3) indicate that the average person who completes 16 and 17 years of schooling will live 9.1 and 20.8 percent longer than high school graduates. Conversely, there seems to a smaller than average relationship between the first eleven years of education and mortality. As shown above, when we use instruments to predict variation in educational attainment that is limited to education in high school or less, there is no causal effect of attained years of schooling on mortality. Our analysis does not allow us to say anything about whether or not a person would live longer or be less likely to die if it were possible to exogenously assign them an extra year of post-secondary education. That work remains for the future.

3.7 Conclusion

In this paper we use IV methods to investigate whether educational attainment up through high school affects how long a person lives past age 40 and the risk of mortality in 5 and 10-year age ranges. Using plausibly exogenous variation from changes in compulsory schooling laws across states and over time within a state, we find no evidence that it does. None of the evidence we generate suggests that a person will live longer if one could exogenously change his years of attained education up to grade 12. On the other hand, our instruments are often weak predictors of education and our estimates from instrumental variables models lack precision. Therefore, the IV results provide little information to support or refute the hypothesis that education causes longer lives.

There are number of usual suspects for our difficulty estimating IV models. For the 5- and 10-year mortality models, we must restrict our samples to relatively small groups that are both at risk to die and we observe. Our instrument is strongest in our hazard models, which pool our entire sample. However, after including state fixed effects the predictive power of our instruments in these models is weak. Independent of sample size, there may not be enough within state variation in compulsory schooling laws to predict education in models that include state fixed effects. Finally, compulsory schooling laws may be weakly enforced and play only a minor role in education decisions.

Setting aside our difficulties estimating models testing causality, there are several reasons that we might find no effect of education on mortality. The most obvious explanation is that the type of learning one acquires in high school does little to change how people behave. We do not have the data to directly test that hypothesis but we present supplementary evidence showing a strong association with higher levels of schooling and length of life. Taken literally, the results provide some account of the costs and benefits of compulsory schooling policies. Many states are considering whether or not to require that youth stay in school until age 18. Our results suggest that, while there may be benefits to such policies, added length of life is not one of them.

The idea that better educated people live longer is intuitively appealing. And it is plausible that it is only in college that a person learns what he needs to more efficiently produce health or to allocate resources more effectively. While this analysis does not provide evidence to test that hypothesis, it provides evidence that suggests that researchers should concentrate their efforts on finding exogenous variation in post-secondary education. Though the non-results are disappointing, they also are important because they help focus future research. The evidence we present here should steer researchers in potentially fruitful directions.

Table 3.1: PSID Main Sample Summary Statistics

	Mean	Std. Dev.
Yrs Ed	11.08	3.43
Years Comp Ed	8.74	1.15
Female	0.51	0.50
Black	0.23	0.42
Hispanic	0.03	0.18
Other Race	0.02	0.13
Decade of Birth - 1900	0.09	0.29
Decade of Birth - 1910	0.18	0.39
Decade of Birth - 1920	0.28	0.45
Decade of Birth - 1930	0.28	0.45
Decade of Birth - 1940	0.17	0.38
Parent's Ed: Less than HS	0.78	0.42
Parent's Ed: Some College	0.04	0.20
Parent's Ed: College Degree or Higher	0.06	0.24
N	6145	

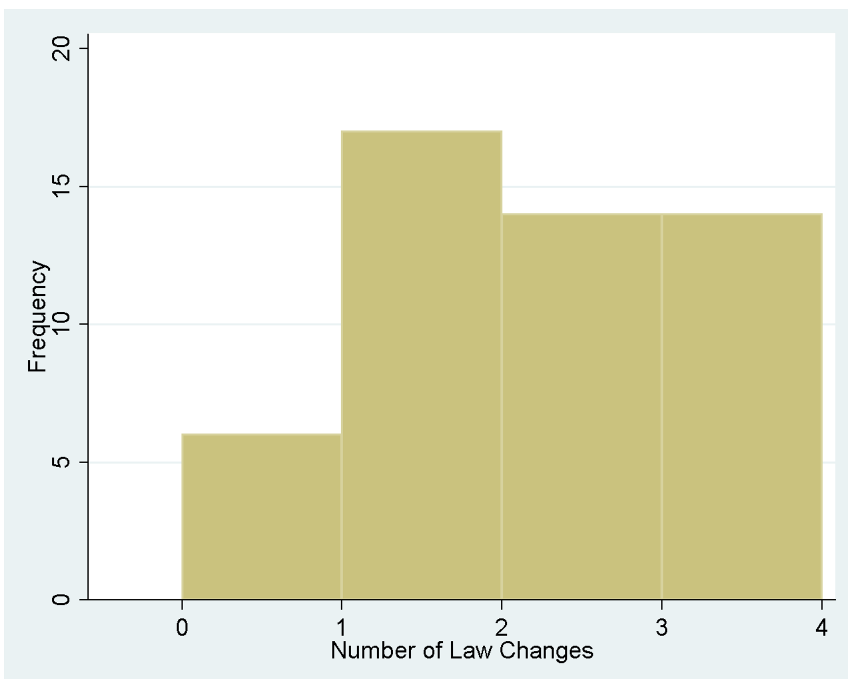

Figure 3.1: Count of State Changes to Compulsory Schooling Laws before 1960.

Table 3.2: First Stage Estimates - Comparison of Cohorts

Cohort	All	1880-1900	1901-1944	1945-1959
Years Comp Ed	0.1956**	-0.3751	0.2204**	-0.1368
	(0.0829)	(0.3021)	(0.0855)	(0.1071)
Parent's Ed: Less than HS	-1.4363***	-2.3504***	-1.7634***	-1.1336***
	(0.1271)	(0.7386)	(0.1493)	(0.1212)
Parent's Ed: Some College	0.7585***	0.1455	1.0897***	0.5502***
	(0.1109)	(1.6102)	(0.1524)	(0.1728)
Parent's Ed: College Degree or Higher	1.6426***	1.2576	1.7332***	1.6707***
	(0.0901)	(1.5964)	(0.1464)	(0.1138)
State Fixed Effects	Yes	Yes	Yes	Yes
N	10667	253	6145	4269
R^2	0.316	0.241	0.255	0.227
F-Statistic on Instruments	5.56	1.54	6.64	1.63
Partial R^2 on Instruments	0.00	0.00	0.00	0.00

Standard errors clustered at state of residence at age 14 in parentheses.
All models include gender, race, and decade of birth controls
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.3: First Stage Estimates

	(1)	(2)	(3)	(4)
Years Comp Ed	0.3005***	0.2982***	0.1898**	0.2204**
	(0.0624)	(0.0585)	(0.0924)	(0.0855)
Parent's Ed: Less than HS		-1.8714***		-1.7634***
		(0.1311)		(0.1493)
Parent's Ed: Some College		1.3120***		1.0897***
		(0.1407)		(0.1524)
Parent's Ed: College Degree or Higher		1.6318***		1.7332***
		(0.1507)		(0.1464)
State Fixed Effects	No	No	Yes	Yes
N	6145	6145	6145	6145
R^2	0.118	0.213	0.169	0.255
F-Statistic on Instruments	23.18	25.99	4.22	6.64
Partial R^2 on Instruments	0.01	0.01	0.00	0.00

Standard errors clustered at state of residence at age 14 in parentheses.
Sample restricted to those born from 1901 to 1945.
All models include gender, race, and decade of birth controls.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

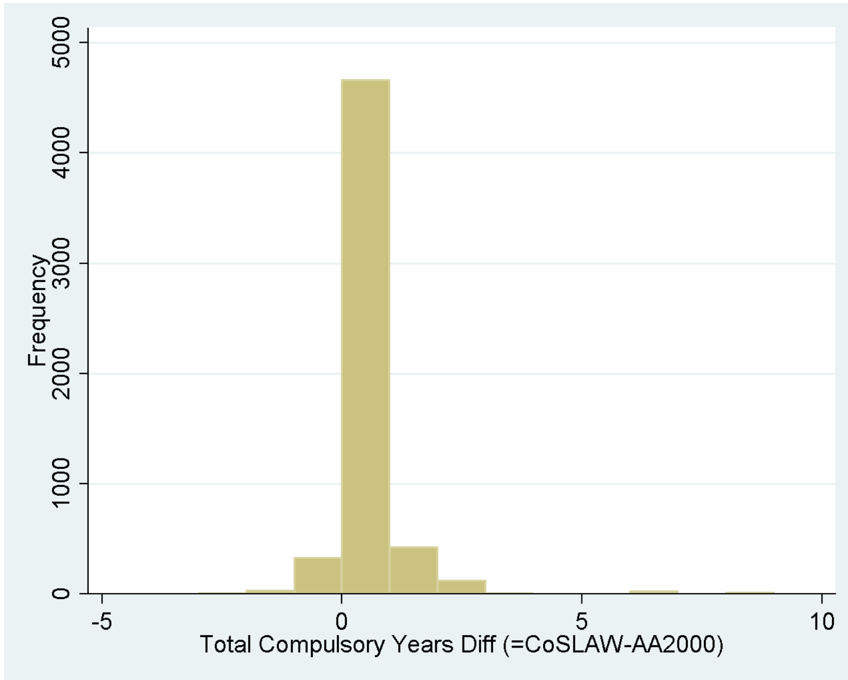

Figure 3.2: Distribution of differences in compulsory years between CoSLAW database and Angrist & Acemoglu (2000).

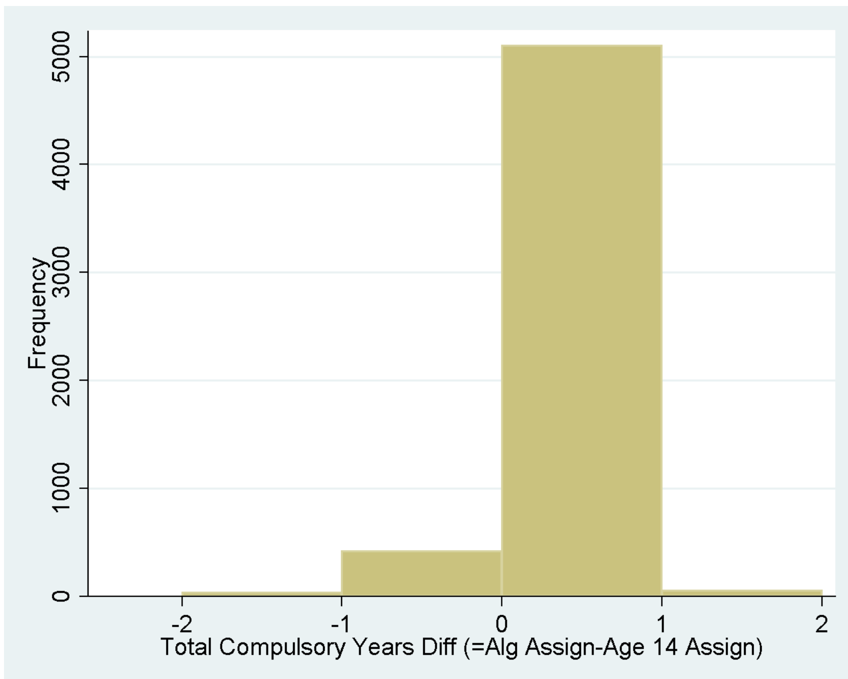

Figure 3.3: Distribution of differences in compulsory years between assignment algorithm and age 14 rule using CoSLAW database.

Table 3.4: First Stage Estimates - Comparison of Instruments

	Angrist & Acemoglu 2000				CoSLAW	
	(1)	(2)	(3)	(4)	(5)	(6)
Comp Att w/ Yrs Req Sch (Age 14)	0.1053**					
	(0.0483)					
Comp Att w/o Yrs Req Sch (Age 14)		0.1523**				
		(0.0623)				
Comp Att w/o Yrs Req Sch (Alg Assign)			0.1037*			
			(0.0563)			
Child Labor Laws (Age 14)				0.0827		
				(0.0577)		
Comp Att (Age 14)					0.1931**	
					(0.0775)	
Comp Att (Alg Assign)						0.2876***
						(0.0836)
Parent's Ed: Less than HS	-1.7327***	-1.7349***	-1.7355***	-1.7321***	-1.7325***	-1.7305***
	(0.1400)	(0.1406)	(0.1400)	(0.1408)	(0.1404)	(0.1400)
Parent's Ed: Some College	1.2027***	1.1966***	1.2009***	1.1974***	1.2077***	1.2133***
	(0.1618)	(0.1621)	(0.1605)	(0.1615)	(0.1639)	(0.1624)
Parent's Ed: College Degree or Higher	1.7749***	1.7771***	1.7816***	1.7774***	1.7791***	1.7860***
	(0.1489)	(0.1496)	(0.1498)	(0.1493)	(0.1505)	(0.1497)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	5615	5615	5615	5615	5615	5615
R^2	0.250	0.250	0.249	0.249	0.250	0.251
F-Statistic on Instruments	4.75	5.98	3.39	2.05	6.20	11.82
Partial R^2 on Instruments	0.00	0.00	0.00	0.00	0.00	0.00

Standard errors clustered at state of residence at age 14 in parentheses. Sample restricted to those born from 1901 to 1945.
All models include gender, race, and decade of birth controls
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: First Stage Estimates - Comparison of Instruments without State Fixed Effects

	Angrist & Acemoglu 2000				CoSLAW	
	(1)	(2)	(3)	(4)	(5)	(6)
Comp Att w/ Yrs Req Sch (Age 14)	0.1847***					
	(0.0511)					
Comp Att w/o Yrs Req Sch (Age 14)		0.2726***				
		(0.0580)				
Comp Att w/o Yrs Req Sch (Alg Assign)			0.2197***			
			(0.0577)			
Child Labor Laws (Age 14)				0.2291***		
				(0.0630)		
Comp Att (Age 14)					0.2693***	
					(0.0608)	
Comp Att (Alg Assign)						0.3106***
						(0.0613)
Parent's Ed: Less than HS	-1.8431***	-1.8458***	-1.8454***	-1.8438***	-1.8360***	-1.8320***
	(0.1183)	(0.1204)	(0.1189)	(0.1207)	(0.1185)	(0.1184)
Parent's Ed: Some College	1.4394***	1.4325***	1.4455***	1.4328***	1.4577***	1.4622***
	(0.1289)	(0.1337)	(0.1316)	(0.1370)	(0.1361)	(0.1355)
Parent's Ed: College Degree or Higher	1.6972***	1.7090***	1.7160***	1.6919***	1.7062***	1.7139***
	(0.1515)	(0.1564)	(0.1572)	(0.1501)	(0.1601)	(0.1598)
State Fixed Effects	No	No	No	No	No	No
N	5615	5615	5615	5615	5615	5615
R^2	0.203	0.205	0.204	0.202	0.204	0.206
F-Statistic on Instruments	13.09	22.11	14.48	13.21	19.63	25.65
Partial R^2 on Instruments	0.01	0.01	0.01	0.01	0.01	0.01

Standard errors clustered at state of residence at age 14 in parentheses. Sample restricted to those born from 1901 to 1945.
All models include gender, race, and decade of birth controls
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: Auxiliary Linear Probability Models of Educational Attainment

	(1) 6 Years	(2) 7 Years	(3) 8 Years	(4) 9 Years	(5) 10 Years	(6) 11 Years	(7) 12 Years	(8) 13 Years	(9) 14 Years	(10) 15 Years	(11) 16 Years
Years Comp Ed	0.0165***	0.0222**	0.0353***	0.0247*	0.0291**	0.0273*	0.0235	0.0092	0.0073	0.0049	0.0034
	(0.0052)	(0.0093)	(0.0112)	(0.0139)	(0.0144)	(0.0140)	(0.0149)	(0.0069)	(0.0058)	(0.0058)	(0.0059)
Parent's Ed: Less than HS	-0.0404***	-0.0779***	-0.0911***	-0.1532***	-0.1773***	-0.2096***	-0.2324***	-0.1976***	-0.1739***	-0.1442***	-0.1252***
	(0.0109)	(0.0220)	(0.0234)	(0.0227)	(0.0204)	(0.0184)	(0.0155)	(0.0132)	(0.0140)	(0.0117)	(0.0120)
Parent's Ed: Some College	0.0244***	0.0270***	0.0407***	0.0626***	0.0867***	0.1128***	0.1157***	0.1740***	0.1566***	0.1104***	0.0985***
	(0.0070)	(0.0072)	(0.0099)	(0.0163)	(0.0178)	(0.0191)	(0.0200)	(0.0270)	(0.0284)	(0.0269)	(0.0289)
Parent's Ed: College Degree or Higher	0.0282***	0.0398***	0.0561***	0.0791***	0.0896***	0.1113***	0.1346***	0.2675***	0.2691***	0.2549***	0.2523***
	(0.0048)	(0.0076)	(0.0106)	(0.0150)	(0.0132)	(0.0146)	(0.0168)	(0.0237)	(0.0269)	(0.0261)	(0.0297)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	6145	6145	6145	6145	6145	6145	6145	6145	6145	6145	6145
R^2	0.080	0.121	0.139	0.170	0.179	0.193	0.205	0.170	0.162	0.147	0.141

Standard errors clustered at state of residence at age 14 in parentheses. Sample restricted to those born from 1901 to 1945.
All models include gender, race, and decade of birth controls
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

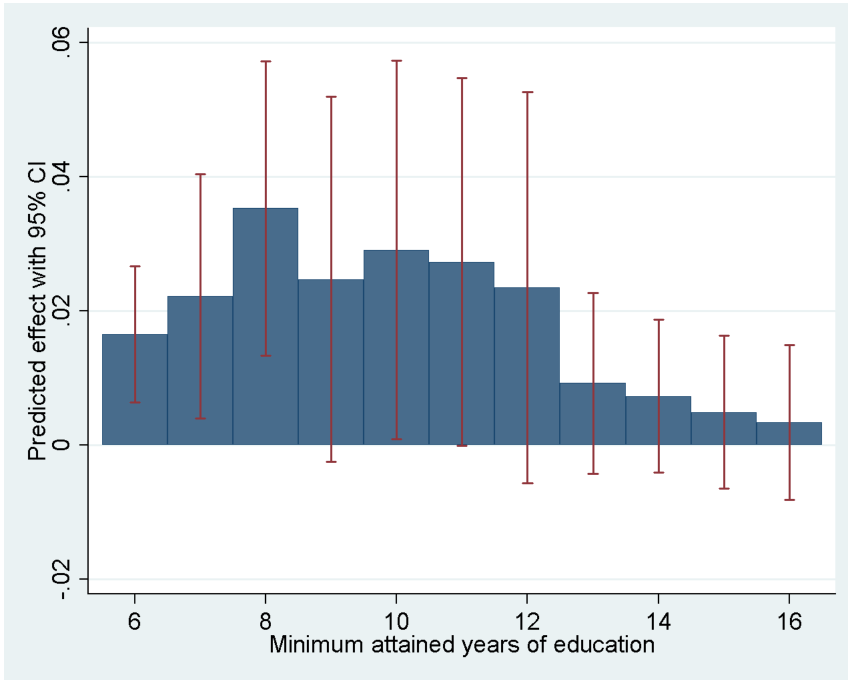

Figure 3.4: Effect of years of required attendance on the probability of completing a minimum level of schooling.

Table 3.7: 10-Year Mortality Rates

	Mean	N
Died Age 50-59	0.10	2784
Died Age 60-69	0.19	2565
Died Age 70-79	0.39	1690
Died Age 80-89	0.64	537
N	3991	

Table 3.8: Ten-Year Mortality: Naive Probit

	(1) Died Age 50-59	(2) Died Age 60-69	(3) Died Age 70-79	(4) Died Age 80-89
Yrs Ed	-0.0052**	-0.0078**	-0.0102**	-0.0191***
	(0.0023)	(0.0033)	(0.0043)	(0.0056)
Parent's Ed: Less than HS	0.0047	0.0133	0.0640*	-0.1070*
	(0.0124)	(0.0206)	(0.0363)	(0.0611)
Parent's Ed: Some College	0.0010	-0.0058	-0.0057	-0.0590
	(0.0219)	(0.0336)	(0.0699)	(0.1162)
Parent's Ed: College Degree or Higher	-0.0521**	-0.0103	0.0716	-0.1579**
	(0.0158)	(0.0434)	(0.0553)	(0.0813)
N	2784	2565	1690	537
Pseudo R^2	0.065	0.043	0.047	0.065

Marginal effects reported. Standard errors clustered at state of residence at age 14 in parentheses.
Sample restricted to those born from 1901 to 1945.
All models include gender, race, and decade of birth controls.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.9: Ten-Year Mortality: 2SRI Probit

	(1) Died Age 50-59	(2) Died Age 60-69	(3) Died Age 70-79	(4) Died Age 80-89
Yrs Ed	-0.0220	0.0359	4.9402**	0.7517
	(0.0305)	(0.0458)	(2.4084)	(1.3982)
Yrs Ed - Residual	0.0168	-0.0439	-4.9504**	-0.7709
	(0.0304)	(0.0458)	(2.4070)	(1.3985)
Parent's Ed: Less than HS	-0.0140	0.0650	1.0000**	0.8635
	(0.0401)	(0.0533)	(0.0001)	(0.3989)
Parent's Ed: Some College	0.0213	-0.0576	-0.5202**	-0.6910
	(0.0521)	(0.0541)	(0.0656)	(0.2021)
Parent's Ed: College Degree or Higher	-0.0356	-0.0786	-0.8652**	-0.7551
	(0.0395)	(0.0654)	(0.1481)	(0.1539)
State Fixed Effects	Yes	Yes	Yes	Yes
N	2784	2565	1690	537
Pseudo R^2	0.065	0.043	0.048	0.065
F-Statistic on Instruments	7.39	8.83	0.00	0.02
Partial R^2 on Instruments	0.00	0.00	0.00	0.00

Marginal effects reported. Standard errors clustered at state of residence at age 14 in parentheses.
Sample restricted to those born from 1901 to 1945.
All models include gender, race, and decade of birth controls.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.10: Ten-Year Mortality: 2SLS (Linear Probability Model)

	(1) Died Age 50-59	(2) Died Age 60-69	(3) Died Age 70-79	(4) Died Age 80-89
Yrs Ed	-0.0265	0.0440	4.7594**	0.7160
	(0.0372)	(0.0510)	(2.3097)	(1.4101)
Yrs Ed - Residual	0.0207	-0.0523	-4.7690**	-0.7334
	(0.0370)	(0.0512)	(2.3084)	(1.4104)
Parent's Ed: Less than HS	-0.0187	0.0787	8.7147**	1.2784
	(0.0442)	(0.0686)	(4.1939)	(2.6476)
Parent's Ed: Some College	0.0243	-0.0730	-3.2050**	-0.9970
	(0.0499)	(0.0719)	(1.5470)	(1.8135)
Parent's Ed: College Degree or Higher	-0.0114	-0.1022	-8.5116**	-1.8844
	(0.0608)	(0.0960)	(4.1521)	(3.3610)
State Fixed Effects	Yes	Yes	Yes	Yes
N	2784	2565	1690	537
R^2	0.042	0.042	0.062	0.081
F-Statistic on Instruments	7.39	8.83	0.00	0.02
Partial R^2 on Instruments	0.00	0.00	0.00	0.00

Linear probability models. Standard errors clustered at state of residence at age 14 in parentheses.
Sample restricted to those born from 1901 to 1945.
All models include gender, race, and decade of birth controls.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.11: 5-Year Mortality Rates

	Mean	N
Died Age 45-49	0.03	2265
Died Age 50-54	0.04	2737
Died Age 55-59	0.06	3113
Died Age 60-64	0.08	3111
Died Age 65-69	0.12	2642
Died Age 70-74	0.18	2084
Died Age 75-79	0.26	1400
Died Age 80-84	0.37	767
N	4804	

Table 3.12: Five-Year Mortality: Naive Probit

	(1) Died Age 45-49	(2) Died Age 50-54	(3) Died Age 55-59	(4) Died Age 60-64	(5) Died Age 65-69	(6) Died Age 70-74	(7) Died Age 75-79	(8) Died Age 80-84
Yrs Ed	-0.0028***	-0.0025**	-0.0024	-0.0026	-0.0060***	-0.0038	-0.0078**	-0.0161***
	(0.0009)	(0.0012)	(0.0016)	(0.0019)	(0.0023)	(0.0025)	(0.0035)	(0.0054)
Parent's Ed: Less than HS	0.0012	0.0031	0.0060	0.0276**	-0.0054	0.0423	0.0053	-0.0789
	(0.0072)	(0.0096)	(0.0100)	(0.0108)	(0.0153)	(0.0344)	(0.0272)	(0.0581)
Parent's Ed: Some College	-0.0125	-0.0020	0.0086	-0.0358	0.0211	-0.0212	-0.0163	-0.0742
	(0.0092)	(0.0199)	(0.0163)	(0.0172)	(0.0331)	(0.0443)	(0.0568)	(0.0911)
Parent's Ed: College Degree or Higher	0.0140	-0.0230	-0.0219	-0.0020	0.0134	0.0323	0.0215	-0.0279
	(0.0184)	(0.0128)	(0.0144)	(0.0275)	(0.0284)	(0.0481)	(0.0369)	(0.0615)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2265	2737	3113	3111	2642	2084	1400	767
Pseudo R^2	0.062	0.028	0.045	0.033	0.020	0.021	0.025	0.030

Marginal effects reported. Standard errors clustered at state of residence at age 14 in parentheses.
Sample restricted to those born from 1901 to 1945.
All models include gender, race, and decade of birth controls.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.13: Five-Year Mortality: 2SRI Probit

	(1) Died Age 45-49	(2) Died Age 50-54	(3) Died Age 55-59	(4) Died Age 60-64	(5) Died Age 65-69	(6) Died Age 70-74	(7) Died Age 75-79	(8) Died Age 80-84
Yrs Ed	0.0187*	-0.0227	-0.0093	-0.0043	0.0371	0.2815	-0.4004	-0.5464**
	(0.0109)	(0.0168)	(0.0095)	(0.0164)	(0.0432)	(0.3235)	(0.3124)	(0.2378)
Yrs Ed - Residual	-0.0219**	0.0205	0.0069	0.0021	-0.0435	-0.2864	0.3929	0.5295**
	(0.0109)	(0.0169)	(0.0099)	(0.0169)	(0.0433)	(0.3239)	(0.3134)	(0.2385)
Parent's Ed: Less than HS	0.0209**	-0.0270	-0.0022	0.0213	0.0459	0.2881	-0.7108	-0.7880**
	(0.0075)	(0.0341)	(0.0153)	(0.0184)	(0.0526)	(0.1744)	(0.4023)	(0.1325)
Parent's Ed: Some College	-0.0187**	0.0196	0.0162	-0.0371	-0.0257	-0.1556	0.3147	0.5137**
	(0.0034)	(0.0356)	(0.0242)	(0.0198)	(0.0477)	(0.0752)	(0.2982)	(0.1632)
Parent's Ed: College Degree or Higher	-0.0120	0.0047	-0.0181	-0.0041	-0.0573	-0.1902	0.7557	0.7064**
	(0.0095)	(0.0454)	(0.0210)	(0.0378)	(0.0500)	(0.0593)	(0.2015)	(0.0517)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2265	2737	3113	3111	2642	2084	1400	767
Pseudo R^2	0.096	0.057	0.060	0.060	0.042	0.042	0.052	0.073
F-Statistic on Instruments	11.65	8.70	12.24	13.03	3.65	0.10	0.14	0.74
Partial R^2 on Instruments	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00

Marginal effects reported. Standard errors clustered at state of residence at age 14 in parentheses.
Sample restricted to those born from 1901 to 1945.
All models include gender, race, and decade of birth controls.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.14: Five-Year Mortality: 2SLS (Linear Probability Model)

	(1) Died Age 45-49	(2) Died Age 50-54	(3) Died Age 55-59	(4) Died Age 60-64	(5) Died Age 65-69	(6) Died Age 70-74	(7) Died Age 75-79	(8) Died Age 80-84
Yrs Ed	0.0330**	-0.0333	-0.0089	0.0003	0.0381	0.2828	-0.4407	-0.4953**
	(0.0143)	(0.0217)	(0.0098)	(0.0185)	(0.0456)	(0.3343)	(0.3151)	(0.2040)
Yrs Ed - Residual	-0.0374**	0.0307	0.0060	-0.0033	-0.0451	-0.2873	0.4334	0.4797**
	(0.0139)	(0.0217)	(0.0104)	(0.0193)	(0.0458)	(0.3347)	(0.3161)	(0.2047)
Parent's Ed: Less than HS	0.0461***	-0.0364	-0.0026	0.0242	0.0510	0.5198	-0.7457	-0.9250**
	(0.0164)	(0.0306)	(0.0150)	(0.0235)	(0.0670)	(0.5681)	(0.5492)	(0.3701)
Parent's Ed: Some College	-0.0492**	0.0271	0.0138	-0.0370	-0.0311	-0.2686	0.2921	0.5028**
	(0.0177)	(0.0280)	(0.0207)	(0.0286)	(0.0597)	(0.3042)	(0.2387)	(0.2177)
Parent's Ed: College Degree or Higher	-0.0394	0.0296	-0.0115	-0.0104	-0.0712	-0.4538	0.8779	1.0579**
	(0.0238)	(0.0390)	(0.0228)	(0.0379)	(0.0826)	(0.5626)	(0.6154)	(0.4785)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2265	2737	3113	3111	2642	2084	1400	767
R^2	0.026	0.021	0.028	0.031	0.031	0.040	0.059	0.091
F-Statistic on Instruments	11.65	8.70	12.24	13.03	3.65	0.10	0.14	0.74
Partial R^2 on Instruments	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.00

Linear probability models. Standard errors clustered at state of residence at age 14 in parentheses.
Sample restricted to those born from 1901 to 1945.
All models include gender, race, and decade of birth controls.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.15: Parametric Survial Model of Mortality After Age 40

	(1)	(2)	(3)	(4)	(5)	(6)
Model	Naive	2SRI	Naive	2SRI	Naive	2SRI
Yrs Ed	0.0137***	-0.0223	0.0124***	-0.0231	0.0135***	-0.1118**
	(0.0027)	(0.0163)	(0.0029)	(0.0163)	(0.0028)	(0.0531)
Residuals		0.0364**		0.0360**		0.1255**
		(0.0163)		(0.0164)		(0.0523)
Parent's Ed: Less than HS			-0.0328	-0.1004***	-0.0314	-0.2538***
			(0.0223)	(0.0343)	(0.0217)	(0.0961)
Parent's Ed: Some College			0.0301	0.0758	0.0345	0.1681**
			(0.0438)	(0.0463)	(0.0467)	(0.0741)
Parent's Ed: College Degree or Higher			-0.0026	0.0529	-0.0094	0.2045**
			(0.0352)	(0.0478)	(0.0361)	(0.0944)
State Fixed Effects	No	No	No	No	Yes	Yes
N	6145	6145	6145	6145	6145	6145
F-Statistic on Instruments		23.18		25.99		6.64
Partial R^2 on Instruments		0.01		0.01		0.00

Standard errors clustered at state of residence at age 14 in parentheses.
Models estimated with Weibull survival distribution in accelerated failure-time metric.
All models include gender, race, and decade of birth controls
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.16: Ten-Year Mortality by Causes Related to Health Behaviors: Naive and 2SRI Probit

	Died Age 50-59		Died Age 60-69		Died Age 70-79		Died Age 80-89	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model	Naive	2SRI	Naive	2SRI	Naive	2SRI	Naive	2SRI
Yrs Ed	-0.0024	0.0185	-0.0042	0.0448	-0.0013	0.1968	-0.0193***	0.1338
	(0.0017)	(0.0215)	(0.0032)	(0.0352)	(0.0049)	(0.2744)	(0.0070)	(0.1663)
Yrs Ed - Residual		-0.0210		-0.0493		-0.1981		-0.1536
		(0.0217)		(0.0358)		(0.2736)		(0.1653)
Parent's Ed: Less than HS	-0.0036	0.0165	0.0201	0.0698*	0.0269	0.2872	-0.1450*	0.1182
	(0.0101)	(0.0205)	(0.0193)	(0.0348)	(0.0309)	(0.2418)	(0.0781)	(0.2887)
Parent's Ed: Some College	0.0009	-0.0179	-0.0026	-0.0488*	-0.0036	-0.1115	-0.1006	-0.2964
	(0.0157)	(0.0171)	(0.0326)	(0.0229)	(0.0603)	(0.1316)	(0.1466)	(0.2085)
Parent's Ed: College Degree or Higher	-0.0141	-0.0296	-0.0145	-0.0753	0.0342	-0.2216	-0.1469	-0.4399
	(0.0128)	(0.0127)	(0.0351)	(0.0323)	(0.0580)	(0.1873)	(0.0981)	(0.2193)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2407	2407	2180	2180	1436	1436	378	378
Pseudo R^2	0.097	0.098	0.048	0.050	0.118	0.119	0.067	0.068
F-Statistic on Instruments		15.41		12.23		0.32		0.50
Partial R^2 on Instruments		0.01		0.01		0.00		0.00

Marginal effects reported. Standard errors clustered at state of residence at age 14 in parentheses.
Sample restricted to those born from 1901 to 1945. All models include gender, race, and decade of birth controls.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.17: Ten-Year Mortality by Causes Unrelated to Health Behaviors: Naive and 2SRI Probit

	Died Age 50-59		Died Age 60-69		Died Age 70-79		Died Age 80-89	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model	Naive	2SRI	Naive	2SRI	Naive	2SRI	Naive	2SRI
Yrs Ed	-0.0021	0.0183	-0.0029	-0.0084	-0.0079**	-0.5085***	-0.0208**	0.7000
	(0.0016)	(0.0150)	(0.0020)	(0.0280)	(0.0033)	(0.1404)	(0.0095)	(1.1819)
Yrs Ed - Residual		-0.0205		0.0056		0.5008***		-0.7208
		(0.0153)		(0.0278)		(0.1404)		(1.1844)
Parent's Ed: Less than HS	0.0061	0.0235	0.0020	-0.0053	0.0297	-0.8636***	-0.1541	0.7327
	(0.0084)	(0.0126)	(0.0131)	(0.0379)	(0.0356)	(0.1065)	(0.1078)	(0.4614)
Parent's Ed: Some College	0.0032	-0.0176	0.0092	0.0160	-0.0626	0.3900**	-0.0205	-0.4736
	(0.0180)	(0.0160)	(0.0275)	(0.0474)	(0.0645)	(0.1754)	(0.1504)	(0.1773)
Parent's Ed: College Degree or Higher	-0.0190	-0.0302*	-0.0277	-0.0202	0.0389	0.8425***	-0.2086*	-0.5232
	(0.0121)	(0.0064)	(0.0206)	(0.0449)	(0.0666)	(0.0314)	(0.1065)	(0.1095)
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2172	2172	2107	2107	1343	1343	349	349
Pseudo R^2	0.067	0.068	0.053	0.053	0.109	0.112	0.085	0.085
F-Statistic on Instruments		14.32		15.69		0.62		0.04
Partial R^2 on Instruments		0.01		0.01		0.00		0.00

Marginal effects reported. Standard errors clustered at state of residence at age 14 in parentheses.
Sample restricted to those born from 1901 to 1945. All models include gender, race, and decade of birth controls.
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.18: Parametric Survial Model of Mortality After Age 40: Education Categories

Model	(1) Naive	(2) Naive	(3) Naive
Yrs Ed - 1 to 6	-0.0111	-0.0084	-0.0180
	(0.0298)	(0.0298)	(0.0310)
Yrs Ed - 7 to 8	-0.0587***	-0.0565***	-0.0616***
	(0.0200)	(0.0201)	(0.0195)
Yrs Ed - 9 to 11	-0.0979***	-0.0960***	-0.0950***
	(0.0249)	(0.0246)	(0.0252)
Yrs Ed - 13 to 15	0.0290	0.0269	0.0234
	(0.0332)	(0.0334)	(0.0342)
Yrs Ed - 16	0.0700**	0.0687*	0.0650*
	(0.0354)	(0.0364)	(0.0373)
Yrs Ed - 17	0.2207***	0.2175***	0.2220***
	(0.0338)	(0.0345)	(0.0343)
Parent's Ed: Less than HS		-0.0207	-0.0202
		(0.0227)	(0.0222)
Parent's Ed: Some College		0.0161	0.0205
		(0.0427)	(0.0453)
Parent's Ed: College Degree or Higher		-0.0239	-0.0281
		(0.0349)	(0.0362)
State Fixed Effects	No	No	Yes
N	6145	6145	6145

Standard errors clustered at state of residence at age 14 in parentheses.
Models estimated with Weibull survival distribution in accelerated failure-time metric.
All models include gender, race, and decade of birth controls
Statistical significance of coefficient estimates: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.8 Appendix A: State of Residence Matching Algorithm

Our state of residence matching algorithm uses all information available in the PSID. These data include an individual's current state of residence when interviewed, state of birth for some respondents, a self-identified state in which heads and wives grew up, and the PSID Relationship File. Each of these data identifies a state in a given year and establishes temporal gaps during which a person moved across state lines at least one time. To shrink the number of years in those gaps, we assign a respondent's state of birth to anyone who was in the household at the time of her birth. We assume parents and older siblings (if they were younger than 18 at the time of birth) were present in the household. With these data we shrink the gap over which known moves occurred and sometimes identify a previously unidentified state of residence. Our algorithm produces a residential history that runs from each person's year of birth to the last year she participated in a survey.

However gaps remain in this history and we develop another algorithm to fill in the remaining gaps. We adopt the following rules: 1) when a person lived in the same state on both sides of a gap in the series, we assume she did not move; 2) we use all information on dates of moves (these data identify any move - even when it is from one house to another in the same city) and assign a cross-state move to have occurred on the date that falls into a gap where we know a move occurred; 3) for the remaining cases we impute the date of the move. To impute the move date, we use age-specific probabilities of moving across state lines. We first generate the probability an individual moves to another state for every age using the 1960 U.S. Census. Then, for each individual, we identify the ages associated with the endpoints of the years over which a move occurred. Using these individual specific endpoints, we compute the probability a move occurred in that interval. Using the above age-specific probability distribution, we identify the age in the range that is the midpoint of the conditional probability distribution. That is, over the individual specific age range, we identify the midpoint age as that age where she is equally likely to have moved in the years before and in the years after that age. We assume each cross-state move occurs on this probabilistic-midpoint date. This procedure fills all remaining gaps.

By taking into account all available information in the PSID and by using external, nationally representative information to probabilistically identify dates of moves in remaining gaps, we generate a more accurate state of residence history. Because it uses more information, the assignment is less prone to bias that is introduced when one assumes a respondent always attends school in her state of birth.

3.9 Appendix B: Compulsory Schooling Law Data (CoSLaw)

This appendix describes the coding rules we used to convert statutes into machine readable data. We also compare and contrast our coding of the compulsory schooling laws with data used in Acemoglu and Angrist (2000). We do so only because the Acemoglu and Angrist (2000) data are used quite widely by researchers and so we want to point out places of disagreement in the two data sources so researchers can be clear about possible sources of differences in empirical results.

We compare discrepancies in the coding of state laws on the age by which a parent must send a child to school and the age until which a child has to remain in school. Hereafter we refer to these laws as the minimum mandated age of school entry and the minimum school leaving age respectively.

Definition of terms

The CoSLaw project uses the ages specified in each states printed statutes as passed and enacted by the state legislatures. Compulsory schooling laws typically identify four ages relevant to behavioral models of schooling. There are two ages that pertain to school entry and two ages that pertain to school leaving. States usually set an age by which a parent must send a child to school and an age at which a parent is permitted to send a child to school. We label these ages as the "mandated entry age" and the "permitted entry age." States also set an age at which a youth is permitted to leave school and an age at which he must leave public school. We similarly label these ages as the "permitted exit age" and the "mandated exit age." Although we know of no study to examine the question, it is unlikely that the mandated-exit-age predicts school leaving ages.

The guiding principle in the reading of these laws was that these ages are those stated by the law without any conditions attached to them. That is, the mandated-entry-age is the youngest age by which a parent must send their child to school when no other conditions were met. The permitted-entry-age is the youngest age a child could be enrolled in school when no other conditions were stated or had to be met. The same principle was used to code the permitted-exit-age and the mandated-exit-age. In the few cases where this coding principle is not used, a note is included in the state level documentation.

In addition to these four basic ages, most states also specify conditions under which parents may delay enrolling a child past the mandated-entry-age and conditions under which a youth may leave school at an age younger than the permitted-exit-age. We code and include in CoSLaw the conditions under which exemptions are granted and, if an age or grade was named, the age or grade that applies when those conditions hold.

Coding rules

In coding the state statutes, certain conventions were adopted in the interpretation of ambiguously worded requirements. States commonly neglect to specify whether the required age range is inclusive or not. We assume that unless explicitly stated, age ranges are not inclusive, so a requirement of "eight to sixteen" is coded to apply to children from their eighth birthday until their sixteenth birthday. This is coded identically to a requirement of "eight to fifteen inclusive".

Comparison of CoSLaw data to Acemoglu and Angrist (2000) data

As noted earlier, social scientists are increasingly using variation in compulsory schooling laws to predict patterns of enrollment and educational attainment. Many researchers use data that were compiled by Joshua Angrist, Alan Krueger, and Daron Acemoglu and used in studies they published (see Angrist and Krueger (1991); Acemoglu and Angrist (2000)). Because so many researchers use these data, we next compare and contrast differences in the coding of the compulsory schooling laws in the years that overlap in the two data sets. Whenever possible we note the source of the discrepancy and the basis for the data included in CoSLaw.

In Table 3.19 (below) discrepancies between CoSLaw data and Acemoglu and Angrist (2000) data are documented. In cases where the cause of the discrepancy is clear, the reason is noted. Mandated age of school entry is indicated by "Min." Permitted school leaving age is indicated by "Max." Discrepancies are shown in bold. Year spans are inclusive.

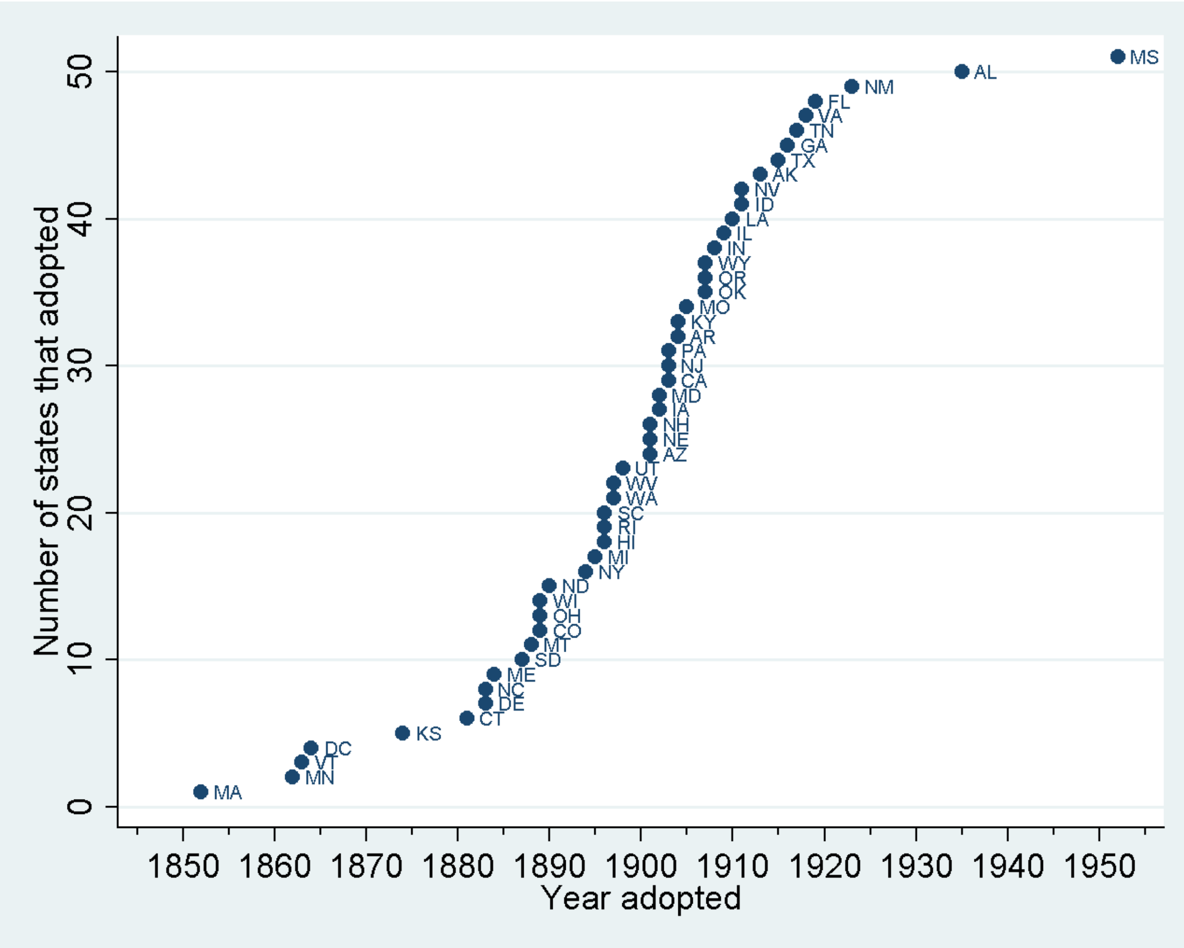

Figure 3.5: Year of adoption of state compulsory schooling laws from CoSLaw data, 1852-1952.

Table 3.19: Conflicts of CoSLaw and A&A2000 data

		CoSLaw		A&A(2000)	
State	Years	Min	Max	Min	Max
AL	1917-1926			**8**	**16**
AL	1927-1934	**7**	16	**8**	16
AR	1917-1928	7	**16**	7	**15**
AR	1978	7	**16**	7	**15**
CA	1914-1918	8	**14**	8	**15**
CA	1919-1920	8	**16**	8	**15**
CA	1967-1977	**6**	16	**8**	16
CO	1963-1964	**7**	16	**8**	16
DE	1924-1952	7	**14**	7	**16**
DE	1969-1971	**6**	16	**7**	16
DC	1925-1934	**7**	**16**	**8**	**14**
FL	1919-1920	**7**	**17**	**8**	**14**
FL	1921-1938	7	**17**	7	**16**
GA	1939-1944	**8**	14	**7**	14
GA	1945	7	**16**	7	**14**
ID	1921-1923	8	**18**	8	**16**
ID	1939-1945	**8**	**18**	**7**	**16**
ID	1949	**7**	**16**	**8**	**18**
IN	1914-1916	7	**16**	7	**14**
IA	1914-1923	7	**17**	7	**16**
KS	1921-1922	8	**15**	8	**16**
KS	1923	7	**16**	8	**16**
KY	1914-1916	7	**17**	7	**12**
KY	1916-1919	7	**17**	7	**16**
LA	1914-1943	**8**	**15**	**7**	**14**
LA	1944-1945	7	**16**	7	**14**
LA	1978	7	**16**	7	**15**
ME	1935-1938	7	**17**	7	**16**

Continued on next page

Table 3.19 – continued from previous page

		CoSLaw		A&A(2000)	
State	Years	Min	Max	Min	Max
ME	1939-1945	7	**17**	7	**14**
ME	1950-1964	7	**17**	7	**16**
ME	1972-1977	7	**17**	7	**16**
MD	1914-1916	**8**	**16**	**7**	**13**
MD	1917-1920	**8**	**16**	**7**	**15**
MD	1921-1923	**8**	16	**7**	16
MD	1978	**7**	16	**6**	16
MA	1978	**7**	16	**6**	16
MI	1944-1945	**6**	16	**7**	16
MN	1954-1978	**8**	16	**7**	16
MS	1921-1923			**7**	**14**
MS	1924-1928	7	**17**	7	**14**
MS	1930-1938	7	**16**	7	**17**
MS	1946-1953	7	**16**	7	**17**
MO	1920	**7**	16	**8**	16
MO	1939-1945	7	**16**	7	**14**
NE	1917-1918	7	**15**	7	**16**
NV	1920-1923	**7**	**18**	**8**	**16**
NV	1956-1958	7	**17**	7	**18**
NH	1914-1916	8	**16**	8	**14**
NH	1953	**6**	16	**8**	16
NM	1921-1922	7	**14**	6	**16**
NM	1923-1945	6	**17**	6	**16**
NM	1978	**6**	17	**8**	17
NY	1914-1916	**8**	16	**7**	16
NY	1916-1920	**7**	16	**8**	16
NY	1968-1971	**6**	16	**7**	16
NC	1914-1916	8	**14**	8	**12**
NC	1923	**7**	14	**8**	14

Continued on next page

Table 3.19 – continued from previous page

		CoSLaw		A&A(2000)	
State	Years	Min	Max	Min	Max
NC	1945	7	**15**	7	**14**
ND	1917-1920	7	**17**	7	**15**
ND	1959-1970	7	**17**	7	**16**
OH	1914-1916	8	**16**	8	**15**
OH	1921-1923	6	**18**	8	**16**
OK	1931-1938	8	**16**	8	**18**
OK	1939-1940	**8**	**16**	**7**	**18**
OR	1921	9	**15**	9	**16**
OR	1922-1928	**8**	16	**9**	16
OR	1929-1934	**8**	**16**	**9**	**18**
OR	1935-1938	**8**	16	**7**	16
OR	1945	**7**	**18**	**8**	**16**
OR	1950-1953	7	**18**	7	**16**
PA	1938	8	**17**	8	**16**
PA	1946-1947	8	**18**	8	**17**
RI	1917-1920	**7**	16	**8**	16
SC					
SD	1915	8	**17**	8	**14**
SD	1916	8	**16**	8	**14**
SD	1924-1930	8	**16**	8	**17**
SD	1939-1945	**7**	**16**	**8**	**17**
TN	1921-1933	**8**	16	**7**	16
TN	1939-1945	**7**	16	**8**	16
TN	1947-1953	7	**17**	7	**16**
TN	1959-1971	7	**17**	7	**16**
TN	1978	7	**17**	7	**16**
TX	1935-1938	**7**	**16**	**8**	**14**
UT	1921-1928	8	**16**	8	**18**
UT	1939-1945	**8**	18	**7**	18

Continued on next page

Table 3.19 – continued from previous page

		CoSLaw		A&A(2000)	
State	Years	Min	Max	Min	Max
UT	1953	**6**	18	**8**	18
VT	1946	**8**	16	**7**	16
VA	1923	8	**14**	8	**12**
VA	1928	**7**	**15**	**8**	14
VA	1929-1934	**7**	15	**8**	15
VA	1968-1971	**6**	**17**	**7**	**16**
VA	1976-1978	**5**	17	**6**	17
WA	1964-1977	8	**18**	8	**16**
WV	1919-1920	**7**	**16**	**8**	**15**
WI	1972-1974	7	**16**	7	**18**
WI	1975-1977	6	**16**	7	**18**
WI	1978	6	**16**	7	**16**
WY	1914-1923	7	**15**	7	**14**
WY	1924-1928	7	**15**	7	**16**
WY	1929-1944	7	**15**	7	**17**
WY	1954-1958	7	**17**	7	**16**
WY	1978	7	**17**	7	**16**

REFERENCES

Acemoglu, D., and J. Angrist. 2000. "How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws." *NBER macroeconomics annual*, 15(1): 9–59.

Adams Beverage Group. 2006*a*. *Adams Beer Handbook.* Norwalk, CT: Adams Beverage Group.

Adams Beverage Group. 2006*b*. *Adams Liquor Handbook.* Norwalk, CT: Adams Beverage Group.

Adams Beverage Group. 2006*c*. *Adams Wine Handbook.* Norwalk, CT: Adams Beverage Group.

Adams, S.J. 2002. "Educational attainment and health: Evidence from a sample of older adults." *Education Economics*, 10(1): 97–109.

Angrist, J.D., and A.B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics*, 979–1014.

Angrist, J.D., and J.S. Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Arendt, J.N. 2005. "Does education cause better health? A panel data analysis using school reforms for identification." *Economics of Education Review*, 24(2): 149–160.

Avery, R.J., D. Kenkel, D. Lillard, and A.D. Mathios. 2007. "Private Profits and Public Health: Does Advertising of Smoking Cessation Products Encourage Smokers to Quit?" *Journal of Political Economy*, 115(3): 447–481.

Bagwell, K. 2007. "The Economic Analysis of Advertising." *Handbook of Industrial Organization*, 3: 1701–1844.

Beer Institute. 2011*a*. "Advertising and Marketing Code." *http://www.beerinstitute.org/tier.asp?bid=249.*

Beer Institute. 2011*b*. "Beer Advertising Facts." *http://www.beerinstitute.org/BeerInstitute/files/ccLibraryFiles/Filename/000000000226/beeradfacts.pdf.*

Berger, M.C., and J.P. Leigh. 1989. "Schooling, self-selection, and health." *Journal of Human Resources*, 433–455.

Black, S.E., P.J. Devereaux, and K. Salvanes. 2004. "Fast times at Ridgemont High? The effect of compulsory schooling laws on teenage births."

Bouchery, E.E., H.J. Harwood, J.J. Sacks, C.J. Simon, and R.D. Brewer. 2011. "Economic Costs of Excessive Alcohol Consumption in the U.S., 2006." *American Journal of Preventive Medicine*, 41(5): 516–524.

Carpenter, C. 2007. "Heavy Alcohol Use and Crime: Evidence from Underage Drunk-Driving Laws." *Journal of Law and Economics*, 50(3): 539–557.

Carpenter, C., and C. Dobkin. 2009. "The Effect of Alcohol Consumption on Mortality: Regression Discontinuity Evidence from the Minimum Drinking Age." *American Economic Journal: Applied Economics*, 1(1): 164–182.

Center on Alcohol Marketing and Youth. 2007. "It Can Be Done: Reductions in Youth Exposure to Alcohol Advertising in Magazines 2001-2005." *Center on Alcohol Marketing and Youth, Georgetown University.*

Center on Alcohol Marketing and Youth. 2010*a*. "Youth Exposure to Alcohol Advertising in National Magazines 2001–2008." *Center on Alcohol Marketing and Youth, Johns Hopkins Bloomberg School of Public Health.*

Center on Alcohol Marketing and Youth. 2010*b*. "Youth Exposure to Alcohol Advertising on Television 2001–2009." *Center on Alcohol Marketing and Youth, Johns Hopkins Bloomberg School of Public Health.*

Centers for Disease Control and Prevention. 2004. "Alcohol-Attributable Deaths and Years of Potential Life Lost - United States, 2001." *Morbidity and Mortality Weekly*, 53(37): 866–870.

Chaloupka, F.J., and K.E. Warner. 2000. "The Economics of Smoking." *Handbook of Health Economics*, 1: 1539–1627.

Chesson, H., P. Harrison, and WJ Kassler. 2000. "Sex Under the Influence: The Effect of Alcohol Policy on Sexually Transmitted Disease Rates in the United States." *Journal of Law and Economics*, 43(1): 215–238.

Chevalier, A. 2004. "Parental education and childs education: A natural experiment." *V IZA Discussion Papers*, , (1153).

Chou, S.Y., J.T. Liu, M. Grossman, and T. Joyce. 2010. "Parental Education and Child Health: Evidence from a Natural Experiment in Taiwan." *American Economic Journal: Applied Economics*, 2(1): 33–61.

Christopoulou, R., J. Han, A. Jaber, and D.R. Lillard. 2011. "Dying for a smoke: How much does differential mortality of smokers affect estimated life-course smoking prevalence?" *Preventive medicine*, 52(1): 66–70.

Clark, D., and H. Royer. 2010. "The effect of education on adult health and mortality: Evidence from Britain." National Bureau of Economic Research.

Collins, R.L., P.L. Ellickson, D. McCaffrey, and K. Hambarsoomians. 2007. "Early Adolescent Exposure to Alcohol Advertising and Its Relationship to Underage Drinking." *Journal of Adolescent Health*, 40(6): 527–534.

Connolly, G.M., S. Casswell, J.I.A.F. Zhang, and P.A. Silva. 1994. "Alcohol in the Mass Media and Drinking by Adolescents: A Longitudinal Study." *Addiction*, 89(10): 1255–1263.

Currie, J., and E. Moretti. 2003. "Mother's education and the intergenerational transmission of human capital: Evidence from college openings." *The Quarterly Journal of Economics*, 118(4): 1495–1532.

Cutler, D.M., and A. Lleras-Muney. 2008. "Education and Health: Evaluating Theories and Evidence." In *Making Americans healthier: social and economic policy as health policy.* , ed. House J.S. Kaplan G. Schoeni, R.F. and H. Pollack, 29. Russell Sage Foundation Publications.

Distilled Spirits Council of the U.S. 2011. "Advertising and Marketing Code." *http://www.discus.org/pdf/May_26_2011_DISCUS_Code_Word_Version.pdf.*

Duffy, M. 1991. "Advertising and the Consumption of Tobacco and Alcoholic Drink: A System-Wide Analysis." *Scottish Journal of Political Economy*, 38(4): 369–385.

Edwards, L.N. 1978. "An empirical analysis of compulsory schooling legislation, 1940-1960." *Journal of Law and Economics*, 21(1): 203–222.

Edwards, L.N., and M. Grossman. 1979. "The Relationship Between Children's Health and Intellectual Development." In *Health, what is it worth?: Measures of health benefits.* , ed. J.H. Stock and D.W.K Andrews. Pergamon Press.

Ellickson, P.L., R.L. Collins, K. Hambarsoomians, and D.F. McCaffrey. 2005. "Does Alcohol Advertising Promote Adolescent Drinking? Results from a Longitudinal Assessment." *Addiction*, 100(2): 235–246.

Färe, R., S. Grosskopf, B.J. Seldon, and V.J. Tremblay. 2004. "Advertising efficiency and the choice of media mix: a case of beer." *International Journal of Industrial Organization*, 22(4): 503–522.

Federal Trade Commission. 1999. "Self-Regulation in the Alcohol Industry: A Review of Industry Efforts to Avoid Promoting Alcohol to Underage Consumers." *http://www.ftc.gov/reports/alcohol/alcoholreport.shtm*.

Federal Trade Commission. 2003. "Alcohol Marketing and Advertising: A Report to Congress." *http://www.ftc.gov/os/2003/09/alcohol08report.pdf*.

Federal Trade Commission. 2008. "Self-Regulation in the Alcohol Industry: Report of the Federal Trade Commission." *http://www.ftc.gov/os/2008/06/080626alcoholreport.pdf*.

Frank, M.W. 2008. "Media Substitution in Advertising: A Spirited Case Study." *International Journal of Industrial Organization*, 26(1): 308–326.

Fuchs, V.R. 1982. "Time Preference and Health: An Exploratory Study." In *Economic Aspects of Health.* , ed. V.R. Fuchs, 93–120. University of Chicago Press for the National Bureau of Economic Research.

Goldin, C., and L. Katz. 2011. "Mass Secondary Schooling and the State: The Role of State Compulsion in the High School Movement." In *Understanding Long-Run Economic Growth.* , ed. D. Costa and N. Lamoreaux, 275–311. University of Chicago Press and National Bureau of Economic Research.

Grossman, M. 1972. "On the concept of health capital and the demand for health." *The Journal of Political Economy*, 80(2): 223–255.

Grossman, M., and R. Kaestner. 1997. "Effects of Education on Health." In *The Social Benefits of Education.* , ed. J.R. Behrman and N. Stacey. University of Michigan Press.

Grossman, Michael. 2006. "Education and Nonmarket Outcomes." In . Vol. 1, , ed. E. Hanushek and F. Welch, 577 – 633. Elsevier.

Hackbarth, D.P., B. Silvestri, and W. Cosper. 1995. "Tobacco and alcohol billboards in 50 Chicago neighborhoods: market segmentation to sell dangerous products to the poor." *Journal of Public Health Policy*, 213–230.

Hastings, Gerard, Susan Anderson, Emma Cooke, and Ross Gordon. 2005. "Alcohol Marketing and Young People's Drinking: A Review of the Research." *Journal of Public Health Policy*, 26(3): 296–311.

Heckman, J.J., F. Flyer, and C. Loughlin. 2008. "An Assessment of Causal Inference in Smoking Initiation Research and a Framework for Future Research." *Economic Inquiry*, 46(1): 37–44.

Hemphill, T.A. 1998. "Harmonizing Alcohol Ads." *Regulation*, 21: 55–62.

Hingson, R., and D. Kenkel. 2004. *National Research Council and Institute of Medicine. Reducing Underage Drinking: A Collective Responsibility.* Washington, DC: National Academies Press.

Hingson, R., T. Heeren, M. Winter, and H. Wechsler. 2005. "Magnitude of Alcohol-Related Mortality and Morbidity Among U.S. College Students Ages 18-24: Changes from 1998 to 2001." *Annual Review of Public Health*, 26: 259–279.

Jenkins, S.P. 2005. "Survival analysis." *Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK. Downloadable from http://www. iser. essex. ac. uk/teaching/degree/stephenj/ec968/pdfs/ec968lnotesv6. pdf.*

Kenkel, D., D. Lillard, and A. Mathios. 2006. "The Roles of High School Completion and GED Receipt in Smoking and Obesity." *Journal of Labor Economics*, 24(3): 635–660.

Kenkel, D.S. 1991. "Health behavior, health knowledge, and schooling." *Journal of Political Economy*, 287–305.

Kenkel, D.S. 1993. "Do Drunk Drivers Pay Their Way? A Note on Optimal Penalties for Drunk Driving." *Journal of Health Economics*, 12(2): 137–149.

Kuo, M., H. Wechsler, P. Greenberg, and H. Lee. 2003. "The marketing of alcohol to college students: the role of low prices and special promotions." *American Journal of Preventive Medicine*, 25(3): 204–211.

Leone, R.P. 1995. "Generalizing what is known about temporal aggregation and advertising carryover." *Marketing Science*, 14(3): 141–150.

Lleras-Muney, A. 2002. "Were Compulsory Attendance and Child Labor Laws Effective? An Analysis from 1915 to 1939." *J. Law & Econ.*, 45: 401–691.

Lleras-Muney, A. 2005. "The relationship between education and adult mortality in the United States." *Review of Economic Studies*, 72(1): 189–221.

Mazumder, B. 2008. "Does education improve health? A reexamination of the evidence from compulsory schooling laws." *Economic Perspectives*, 33(2): 1–15.

McCrary, J., and H. Royer. 2011. "The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth." *American Economic Review*, 101(1).

Milyo, J., and J. Waldfogel. 1999. "The effect of price advertising on prices: Evidence in the wake of 44 liquormart." *American Economic Review*, 1081–1096.

Mokdad, A.H., J.S. Marks, D.F. Stroup, and J.L. Gerberding. 2004. "Actual Causes of Death in the United States, 2000." *Journal of the American Medical Association.*

Molloy, Eamon. 2012. "This Ad's for You: Targeting and the Effect of Alcohol Advertising on Youth Drinking." *Working Paper.*

Mosher, J.F., and D. Johnsson. 2005. "Flavored Alcoholic Beverages: An International Marketing Campaign That Targets Youth." *Journal of Public Health Policy*, 26(3): 326–342.

Nelson, J.P. 1999. "Broadcast Advertising and U.S. Demand for Alcoholic Beverages." *Southern Economic Journal*, 65: 774–790.

Nelson, J.P. 2003. "Advertising Bans, Monopoly, and Alcohol Demand: Testing for Substitution Effects using state panel data." *Review of Industrial Organization*, 22(1): 1–25.

Nelson, J.P. 2010. "What is Learned from Longitudinal Studies of Advertising and Youth Drinking and Smoking? A Critical Assessment." *International Journal of Environmental Research and Public Health*, 7(3): 870.

Nelson, P. 1970. "Information and Consumer Behavior." *Journal of Political Economy*, 78(2): 311–329.

Nelson, P. 1974. "Advertising as Information." *Journal of Political Economy*, 82(4): 729–754.

Rivers, D., and Q.H. Vuong. 1988. "Limited information estimators and exogeneity tests for simultaneous probit models." *Journal of Econometrics*, 39(3): 347–366.

Roberts, M.J., and L. Samuelson. 1988. "An empirical analysis of dynamic, nonprice competition in an oligopolistic industry." *The RAND Journal of Economics*, 200–220.

Saffer, H. 1991. "Alcohol Advertising Bans and Alcohol Abuse: An International Perspective." *Journal of Health Economics*, 10(1): 65–79.

Saffer, H. 1997. "Alcohol Advertising and Motor Vehicle Fatalities." *Review of Economics and Statistics*, 79(3): 431–442.

Saffer, H., and D. Dave. 2002. "Alcohol Consumption and Alcohol Advertising Bans." *Applied Economics*, 34(11): 1325–1334.

Saffer, H., and D. Dave. 2006. "Alcohol Advertising and Alcohol Consumption by Adolescents." *Health Economics*, 15(6): 617–637.

Seldon, B.J., and C. Jung. 1993. "Derived demand for advertising messages and substitutability among the media." *The Quarterly Review of Economics and Finance*, 33(1): 71–86.

Seldon, B.J., R.T. Jewell, and D.M. OBrien. 2000. "Media substitution and economies of scale in advertising." *International Journal of Industrial Organization*, 18(8): 1153–1180.

Snowdon, D.A., S.K. Ostwald, and R.L. Kane. 1989. "Education, survival, and independence in elderly Catholic sisters, 1936-1988." *American Journal of Epidemiology*, 130(5): 999–1012.

Snyder, L.B., F.F. Milici, M. Slater, H. Sun, and Y. Strizhakova. 2006. "Effects of Alcohol Advertising Exposure on Drinking Among Youth." *Archives of Pediatrics and Adolescent Medicine*, 160(1): 18–24.

Stacy, A.W., J.B. Zogg, J.B. Unger, and C.W. Dent. 2004. "Exposure to Televised Alcohol Ads and Subsequent Adolescent Alcohol Use." *American Journal of Health Behavior*, 28(6): 498–509.

Stigler, G.J., and G.S. Becker. 1977. "De Gustibus Non Est Disputandum." *American Economic Review*, 67: 76–90.

Stock, J.H., and M. Yogo. 2001. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg.* , ed. J.H. Stock and D.W.K Andrews, Chapter 5. Cambridge University Press.

Stoops, N. 2004. "Educational Attainment in the United States: 2003. Population Characteristics." *US Department of Commerce*, 12.

Substance Abuse and Mental Health Services Administration. 2007. "Results from the 2006 National Survey on Drug Use and Health: National Findings." *Substance Abuse and Mental Health Services Administration, Office of Applied Studies, Rockville, MD.*

Terza, J.V., A. Basu, and P.J. Rathouz. 2008. "Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling." *Journal of Health Economics*, 27(3): 531–543.

Wine Institute. 2011. "Advertising and Marketing Code." *http://www.wineinstitute.org/initiatives/issuesandpolicy/adcode/details.*

Wyllie, A., J.F. Zhang, and S. Casswell. 1998*a*. "Positive Responses to Televised Beer Advertisements Associated with Drinking and Problems Reported by 18 to 29-Year-Olds." *Addiction*, 93(5): 749–760.

Wyllie, A., J.F. Zhang, and S. Casswell. 1998*b*. "Responses to Televised Alcohol Advertisements Associated with Drinking Behaviour of 10-17-Year-Olds." *Addiction*, 93(3): 361–371.